NATIONAL
URANIUM
TAILINGS
PROGRAM

Canada

A Study to Develop and Demonstrate Techniques for the Probabilistic Assessment of the Long-term Effects of Uranium Mill Tailings in Canada - Phase I

DSS File #26SQ.23241-3-1654
Contract No. OSQ83-00236

A Research Report prepared for:

The National Uranium Tailings Program
Canada Centre for Mineral and Energy Technology
Energy, Mines and Resources Canada
555 Booth Street
Ottawa, Ontario
K1A 0G1

by

SENES Consultants Limited
499 McNicoll Avenue
Willowdale, Ontario
M2H 2C9

MICROMEDIA

251 pp

October 1984

FORWARD

This report is the first of a series in a program of systems model development for the National Uranium Tailings Program. The component modelling in the early phases of the work is being deliberately kept simple (on instructions from the NUTP office) so that attention can focus on the methodology and on the technical meaning, interpretation and use of the results. Thus, specific component models and numerical values should be treated with caution since they are illustrative only. The overall approach taken, what it has to offer, and the possible difficulties with it are far more important than detailed models at this stage. Any comments or suggestions would be welcome.

W. C. Harrison
Project Manager, Modelling
National Uranium Tailings Program
Canada Centre for Mineral and Energy Technology
Energy, Mines and Resources Canada
555 Booth Street
Ottawa, Ontario
K1A 0G1

TABLE OF CONTENTS

		Page
EXECUTIVE SUMMARY		1
1.0 INTRODUCTION		
1.1	Background to the Study	1-1
1.2	Probabilistic Assessment Concept	1-2
1.3	Study Objectives	1-4
1.4	Report Structure	1-5
2.0 APPROACH TO PATHWAYS ANALYSIS		
2.1	Background Considerations	2-1
2.2	Characteristics of Reference Tailings Site	2-3
2.3	Environmental Pathways of Exposure to Man	2-5
2.4	Reference Receptor	2-8
3.0 SYSTEMS CODE AND COMPONENT MODELS		
3.1	Introduction	3-1
3.1.1	Historical Perspectives	3-1
3.1.2	Selection of Appropriate Systems Code	3-4
3.2	Environmental Pathways Descriptions	3-6
3.2.1	Radium in Air	3-7
3.2.2	Radium in Water	3-8
3.2.3	Radon Pathway	3-9
3.2.4	Parameter Value Selection	3-9
3.3	Quality Assurance: Concepts and Applications	3-11
3.3.1	OPANAL QA	3-11
3.3.2	Component Model QA	3-14
4.0 PROBABILISTIC ASSESSMENT RESULTS		
4.1	Statistical Concepts for Interpreting Probabilistic Assessment Results	4-1
4.2	Probabilistic Assessment Results	4-4
4.3	Sensitivity and Uncertainty Analysis	4-7
4.3.1	Sensitivity Analysis	4-8
4.3.2	Uncertainty Analysis	4-10
4.3.3	Combined Sensitivity/Uncertainty Analysis	4-11
4.4	Consideration of Transient Phenomena	4-14
4.5	Dose vs. Risk Assessment	4-16
5.0 CONCLUSIONS AND RECOMMENDATIONS		
5.1	Highlights of Phase I Assessment	5-1
5.2	Recommendations for Further Work in Phase II	5-4
6.0 REFERENCES		

Table of Contents continued

APPENDICES

A SOME USEFUL STATISTICAL CONCEPTS

B OPANAL: A BRIEF DESCRIPTION

C COMPONENT MODELS FOR RADON-222 AND RADIUM-226

D PARAMETER DISTRIBUTIONS

E DISCUSSION PAPERS

E.1 Probabilistic Endpoint for an Assessment
E.2 Sensitivity Analysis for Environmental Simulation Models
E.3 Parameter Uncertainty Analysis Application
E.4 Alternative Sampling Procedures
E.5 Software Quality Assurance: Current Concepts

F STUDY TEAM MEMBERS

A STUDY TO DEVELOP AND DEMONSTRATE TECHNIQUES FOR THE PROBABILISTIC ASSESSMENT OF THE LONG-TERM EFFECTS OF URANIUM MILL TAILINGS IN CANADA - PHASE I

EXECUTIVE SUMMARY

This study, carried out as part of the National Uranium Tailings Program, was conducted to demonstrate the applicability and feasibility of using a probabilistic assessment technique to evaluate the long-term environmental effects of uranium mill tailings. Such an evaluation requires the use of mathematical models to simulate the behaviour of the actual system. The technique has the advantage over traditional deterministic methods, however, in that it incorporates the natural variability and uncertainty regarding model input parameters and provides an opportunity to consider the likelihood of future events directly in the assessment producing a distribution of output results. This Phase I study focussed on examination of the concepts and techniques of probabilistic assessment, and involved a preliminary analysis of a hypothetical "reference" tailings site.

A review of the literature indicated that the probabilistic assessment concept is far from a new idea. Early applications in nuclear research and applied mathematics have been followed by new developments in social science, business and engineering. Recently, probabilistic assessment has been used successfully in a wide variety of environmental and radioecological problems.

The review of probabilistic assessment applications indicated that a variety of computer programs are currently available for this type of analysis. One of these, the OPANAL code, was found to be relatively simple, and easy to implement and was selected for the initial demonstration of the probabilistic approach.

The demonstration of the concept also required a set of mathematical "component" models to describe the source release mechanisms and the environmental effects of radioactive constituents at a hypothetical uranium tailings site. A restricted, simple, yet representative scenario was defined

and the appropriate environmental pathways models were formulated for six key pathways of exposure to a representative receptor. The scenario and models used were kept as simple as possible consistent with the preliminary nature of the study. Despite the simplifications, it is believed that the models exercised all salient features of the probabilistic assessment technique.

The results produced by the probabilistic evaluation of the models using the OPANAL program showed trends common to environmental assessments and illustrated the flexibility of the approach. They indicated that probabilistic assessment could incorporate simple time dependency and also evaluate the effects of different tailings management options. The result of primary interest in this study, however, was considered to be the distribution of radiation dose to a specified individual (i.e. representative receptor). The distribution of the predicted results showed a slightly skewed (lognormal) shape, common in radioecology.

In conclusion, the appropriateness of the probabilistic assessment technique to the evaluation of the long-term environmental effects of uranium mill tailings was demonstrated. Results produced using this technique can be used to provide many different measures of model sensitivity and uncertainty. The results can also be analysed and interpreted in a variety of ways depending on the ultimate "endpoint" of the assessment. The technique provides considerably more information than the traditional deterministic approach about the environmental consequences of the system of interest and thus should contribute to more enlightened decision making.

1.0 INTRODUCTION

1.1 Background to the Study

Considerable work on the behaviour of Canadian uranium mill tailings has been done by many different organizations with different needs or questions in mind at the time the work was performed. This work was reviewed in 1981/82 and Energy, Mines and Resources Canada subsequently established a National Uranium Tailings Program (NUTP) attached to CANMET. The NUTP is a research program intended to pull together the previous work on Canadian uranium mill tailings, to fund additional work where significant gaps are found in understanding or in essential data and to develop the capability to predict the long-term effects of tailings with a known degree of confidence while incorporating natural variability in data and uncertainty in future events. The ability to make quantitative predictions of future behaviour and effects is essential to management, regulatory and public decision making as to the efficacy and acceptability of long-term disposal strategies.

Mathematical modelling plays a major role in this program since quantitative predictions of future behaviour can only be made with a model. The NUTP is, therefore, sponsoring modelling work on two technical levels, one dealing with the various components such as source chemistry, surface water transport or biological uptake and the other providing a framework within which simplified component models can be linked together to model the overall system. This report deals with the first phase of work on modelling the overall system.

The stated objectives of Phase I, as stipulated in the statement of work of the contract were:

"(a) to develop and demonstrate appropriate analysis techniques,
(b) to provide preliminary order-of-magnitude estimates of effects and their uncertainties and,

(c) through a sensitivity analysis, indicate the relative importance of the various components in the system and, thus, identify quantitatively important areas needing further work."

In brief, the Phase I study was to demonstrate the potential usefulness of applying probabilistic assessment techniques to uranium mill tailings and to provide a "first-cut" probabilistic assessment of a (generic) reference tailings site.

1.2 Probabilistic Assessment Concept

Evaluating the potential long-term effects of uranium mill tailings requires mathematical models which simulate the environmental transport and exposure pathways of the various uranium tailings constituents. However, by their very nature, models are only partial representations of real systems and contain a number of parameters which are either only imperfectly known or inherently variable in nature. Consequently, the results of all such calculations are uncertain to some degree.

In general terms there are three main sources of uncertainty in pathways modelling:

- incomplete or invalid models (based on poor understanding or simplifying assumptions)
- inherent parameter bias or variability
- uncertainty regarding future events

Since models are only partial (i.e. imperfect) representations of real systems, a variety of alternative models of varying complexity can often be developed for a particular system. Which (if any) of the models offers the most appropriate representation of reality is often difficult to determine. Thus, regardless of model complexity, there is always an inherent uncertainty associated with the model itself. Indeed increasing the complexity of a model will not necessarily improve its predictive ability, as more complex models require increasing numbers of model parameters which themselves must be

characterized, and often are less easily determined.

Having decided on a model, the information necessary to define the parameters sufficiently may not be available. In other cases the parameters themselves may be inherently stochastic.

In the past, in part because of uncertainties in model formulation and variability in model parameters, there has been a tendency in applying deterministic models to use conservative assumptions to ensure that the calculations overestimate the effect of any release to the environment. In such cases, bias in the selection of parameter values has a substantial effect on the overall degree of conservatism in the final prediction. If many such parameters occur in series and are more or less independent, the cumulative conservatism can be many orders of magnitude. This "overconservatism" often distorts one's perception of reality and thus complicates rather than simplifies decision making. This may not be a problem when predicted effects are so small as to be "negligible" but may become a very real problem when the predicted effects are perceived to be "significant" by either those affected or the regulatory authority or both.

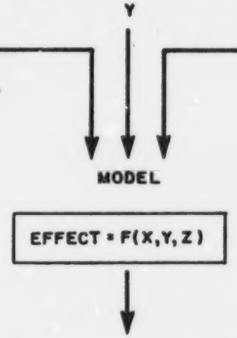
Uncertainties also arise when models are used to predict effects over conditions and periods of time different from those for which the models and model parameters were developed. A further uncertainty in modelling future events arises due to the inability to predict the future sequence and magnitude of rare events which have potential to alter the situation the model is attempting to describe. The inability to predict (with confidence) population distribution in the future is an example of the former type of uncertainty. The inability to predict when a potentially disruptive event (such as an earthquake) might occur is an example of the latter type of uncertainty.

The traditional, deterministic, approach to evaluating environmental effects is illustrated on Figure 1.1. In this approach a single value is selected for each of the model parameters and a single value (the effect) is calculated in

PROBABILISTIC ANALYSIS

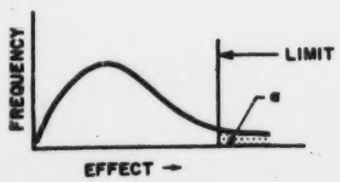
- DEVELOP PARAMETER DISTRIBUTIONS

- INPUT DISTRIBUTIONS TO MODEL



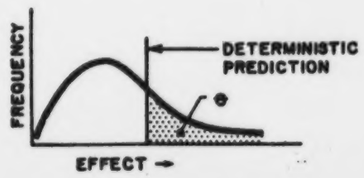
- PRODUCE A DISTRIBUTION OF MODEL PREDICTIONS

- COMPARE WITH LIMIT



α = PROBABILITY OF LIMIT BEING EXCEEDED

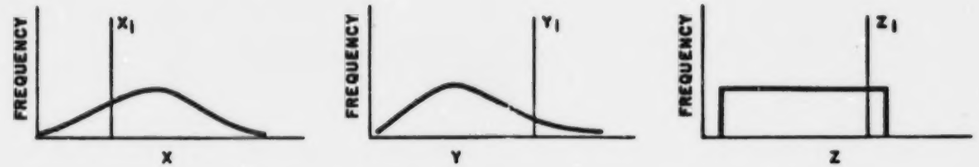
- COMPARE DETERMINISTIC AND PROBABILISTIC RESULTS



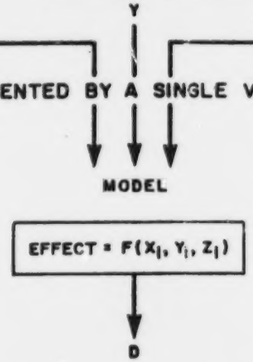
Θ = PROBABILITY OF DETERMINISTIC PREDICTION BEING EXCEEDED

DETERMINISTIC ANALYSIS

- A RANGE OF INPUT VALUES



- PARAMETERS ARE REPRESENTED BY A SINGLE VALUE



- A SINGLE VALUE IS PRODUCED AS MODEL OUTPUT

- COMPARE WITH LIMIT

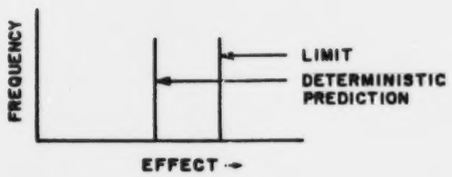


FIGURE 1-1

COMPARISON OF DETERMINISTIC AND PROBABILISTIC ANALYSES

some suitable unit such as concentration, dose or risk. As suggested on the figure, conservative values of the parameters, typically near the upper or lower extreme of the expected range of the frequency distributions for the parameters, are often selected. The influence of the parameter value is generally a very conservative estimate of the "effect" as demonstrated by the following example of a simplified three parameter model of the form:

$$\text{EFFECT} = \frac{Y*Z}{X}$$

If the values of the parameters selected for input to a deterministic analysis are assumed to be varied by a factor of 1.5 from the nominal (i.e. central) values of the parameter distributions, that is increased by a factor of 1.5 in the case of parameters Y and Z and lowered by a factor of 1.5 in the case of parameter X, then the calculated result will be 3.375 times the effect estimated using the nominal parameter values. Similarly, if input parameter values are selected which are 2 to 3 times the nominal parameter values (i.e. the selected parameter values more closely approximate those which would typically be used in a deterministic analysis), then the "effect" would be overestimated in this simple example by a factor of from 8 to 27 times the "effect" estimated using nominal values. With more complex models, the influence of parameter value selection on the deterministic result (i.e. the calculated "effect") can be even more profound.

Calculation of effect using a probabilistic approach follows the same basic route as the deterministic analysis (see Figure 1.1) however, the approach attempts to explicitly account for uncertainty in many of the factors involved. This is achieved by specifying the characteristics of the frequency distributions for each parameter as input to the probabilistic systems code. The model is then run many times with new values specified for the input parameters (X,Y,Z) on each run. The parameter values used in each run (sometimes called a trial) are randomly selected using a random sampling technique which forms part of the code. With this approach, as many individual estimates of the "effect" are obtained as desired by simply increasing or decreasing the number of runs (trials). The results of such a

probabilistic analysis are conveniently summarized by ordering the values into a frequency plot and calculating whatever summary statistics the user desires (eg. mean, quantiles, etc.). The concept of probabilistic analysis, whose outcome is a distribution of effects derived from distributions of input values, is also illustrated in Figure 1.1. The concept has the advantage over the deterministic approach in that the full range of uncertainty in the parameters values may be incorporated into the analysis thus avoiding the usually one-sided bias in the input specifications.

The implications of the results obtained using the deterministic and probabilistic approaches are suggested by the graphical comparison presented at the bottom Figure 1.1. (The concepts involved in carrying out and interpreting a probabilistic analysis are discussed in more detail elsewhere in this report, especially Section 4.1 and Appendix Sections A.11 and A.12.) As indicated previously, the deterministic prediction is generally expected to overestimate the "true" mean effect, whether it be concentration, dose or risk. If the predicted point estimate is less than some limit (as indicated on the figure), then there may be no further need to evaluate the particular design or scenario under consideration. Should the deterministic result exceed the limit however, a complete re-evaluation may be required. The frequency plot of the probabilistic trials on the other hand, shows the full range of possible results. In contrast to the deterministic prediction, it is usually found that a large fraction of the possible probabilistic results are substantially less than the point estimate of effect determined from a deterministic analysis. The probabilistic approach provides the user with considerably more information thus permitting more informed discussion and decision making.

1.3 Study Objectives

Since this was a preliminary study, a number of conceptual aspects of the methodology had to be examined, along with the more practical, technical issues. Several of the conceptual and technical matters, involving questions both of methodology and interpretation of results were addressed in separate

topical papers. To demonstrate some of the more practical aspects, it was agreed to undertake the 'first-cut' probabilistic assessment for a reference tailings site (i.e. generic case study).

The technical objectives were addressed via: the examination of several existing probabilistic assessment tools; adoption and validation of the satisfactory performance of one of the techniques; development of component and transfer models for estimating human exposure from the release of radium-226 and radon-222 via air and water pathways; estimation of parameter values and their associated statistical distributions for a generic reference site; and, application of the probabilistic assessment technique on six selected environmental pathways.

The results of the preliminary evaluation demonstrated the potential usefulness of the technique for estimating the magnitude and uncertainty of potential effects. A preliminary sensitivity analysis was also carried out to demonstrate the type of information which could be gained when this type of analysis is applied to probabilistic results.

1.4 Report Structure

In developing the structure for this report, it was established (in consultation with the Scientific Authority) that the body of the report should avoid excessive detail and present only the basics of the probabilistic assessment methodology and its application to a reference site. Hence, the report was designed to present an abbreviated discussion of: the approach to pathways analysis (Section 2.0); the systems code and component models adopted to describe the reference site (Section 3.0); the results of the preliminary probabilistic assessment (Section 4.0); and recommendations for future work in the next phase of the project (Section 5.0). Brief descriptions of some useful statistical concepts, the OPANAL code, the component models developed to characterize radionuclide movement via the air and water pathways and the values and probability distributions assigned to the model parameters, are included in Appendices A through D. Separate discussion papers on several

key conceptual and technical issues were prepared and these are presented in Appendices E.1 through E.5. Members of the study team are listed in Appendix F.

2.0 APPROACH TO PATHWAYS ANALYSIS

2.1 Background Considerations

Environmental concerns associated with the long-term management of uranium mill tailings focus mainly on the possible effects that radioactivity released from tailings areas might have on man. Radionuclides that can be released from uranium tailings include radon gas from the tailings surface and other radionuclides in windborne dust particles and in groundwater and surface water discharges. Other chemical constituents are also present in the tailings and may be released under certain conditions. The importance of the various release mechanisms is a function of many factors, including, for example, the tailings disposal method employed, the tailings characteristics and meteorological conditions. The concerns apply both during the operating life of uranium mines and after close out, when operations have ceased.

Tailings at uranium mines have generally been directed to one of three basic types of tailings management areas:

- land/shallow-water sites
- dry land sites
- underwater disposal sites

A land/shallow-water site typically features a dam constructed across the outlet of a valley, shallow lake, marsh or low-lying area to provide a tailings containment basin. This is the most common tailings management area employed in Canada.

Dry land containment sites are usually located on flat or gently sloping areas and feature long, low dams or dykes. Such sites are common in the United States.

Underwater disposal refers to the placement of tailings below the water surface, such as at the bottom of a lake or in the ocean. This tailings management method has been used but is much less common than the land/shallow-

water technique.

Assessment of the environmental effects of radioactive releases from uranium tailings management sites has been attempted by a technique known as pathways analysis. Typically, such an analysis is used to estimate the exposure that certain critical individuals or groups (i.e. those most likely to receive the highest exposure) could receive by all of the possible pathways of radiation exposure. A pathways analysis includes assessment of the mechanisms or processes of radionuclide release from tailings, the movement or transportation of radionuclides in the environment (usually via air or water), the uptake or bioaccumulation of radionuclides by local plants and animals, and the living conditions or lifestyle (e.g. eating habits) of the exposed individuals (usually referred to as receptors).

Pathways analyses of uranium mill tailings sites (or indeed uranium mine developments in their entirety) use mathematical models that simulate the behaviour of radionuclides in the environment. Various models have been developed for pathways analysis. It is important to recognize however, that all such calculations contain inherent uncertainties. For example, the models, such as those that predict air dispersion of gaseous or suspended particulate matter are, at best, only approximations of the real world. Moreover, the parameters used in the models are often poorly defined. To account for such uncertainties, there is a tendency to incorporate conservatism into the models which leads to overestimation of exposure or dose. The probabilistic assessment technique has been developed and adopted in several disciplines to avoid introducing undue bias in model predictions. The advantage of such an approach is that the natural variation or uncertainty in parameters (and the occurrence of uncertain events) can be accounted for explicitly.

A complicated example is not necessary (and often undesirable) to demonstrate an idea. Therefore, to demonstrate the applicability of the probabilistic assessment technique, several basic simplifications were agreed upon early in the study, in consultation with the Scientific Authority. These

simplifications limited the number of considerations and clearly defined those factors likely to be of prime importance. It was agreed that the assessment would be carried out for acid-producing tailings placed in a hypothetical reference site. Further, the analysis would be limited to radon-222 and radium-226 (hereafter referred to as radon and radium) and only six key pathways would be considered. The receptor was defined to be an individual living near the reference (generic) site. Finally, dose to the receptor was calculated for assumed quasi-steady state conditions which were defined to extend over a period from a few years to perhaps 200 years following tailings placement. This provided a basic problem which allowed the study team to focus on the techniques yet retain sufficient complexity and realism for the study to be of practical value. A more detailed discussion of these factors is presented in the following sections.

2.2 Characteristics of the Reference Tailings Site

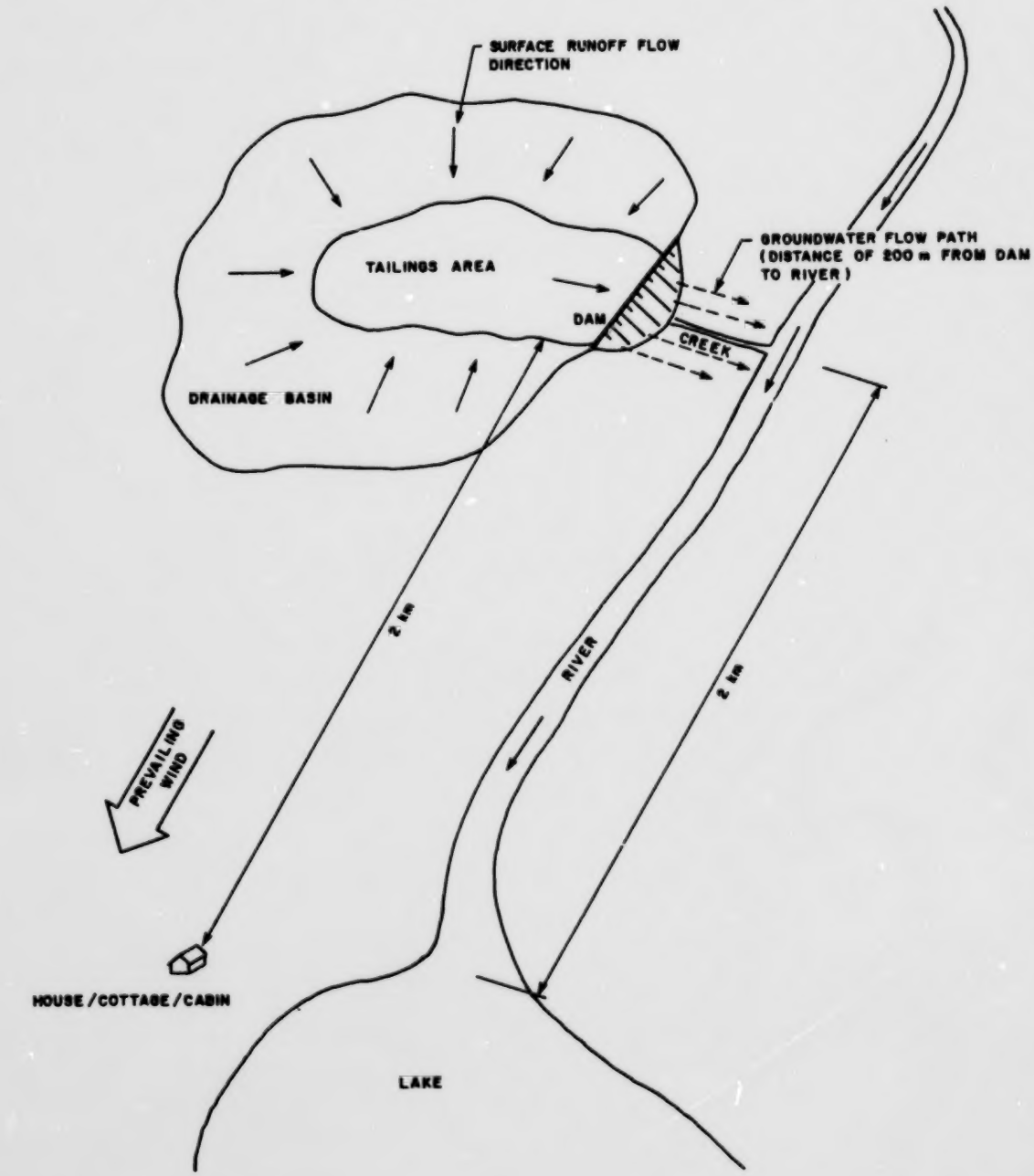
A hypothetical reference tailings site was chosen to reflect the major environmental features and environmental pathways pertinent to tailings management on the Canadian Precambrian Shield. The basic characteristics of the reference site are shown on Figure 2.1. To simplify the analysis, various complexities not relevant to the demonstration of the feasibility of the probabilistic assessment technique were eliminated.

The tailings disposal site was taken to occupy a natural depression formed within the upper reach of a valley, that is, a land/shallow-water type site. The tailings were considered to be contained behind a conventional dam constructed across the low point. The reference tailings site was defined to have a surface area of 50 ha and an average tailings depth of 10 m, for a total volume of 5×10^6 m^3. The drainage basin containing the tailings was assumed to be 150 ha in area (including the tailings.)

The reference tailings characteristics were assumed to represent those resulting from the acid leaching of low grade pyritic uranium ore containing

FIGURE 2·1

SCHEMATIC ILLUSTRATION OF THE REFERENCE TAILINGS SITE



TAILINGS AREA CHARACTERISTICS:

- TAILINGS AREA SIZE - 50 ha
- DRAINAGE BASIN SIZE - 150 ha
- AVERAGE DEPTH OF TAILINGS - 10 m
- VOLUME OF TAILINGS - 5×10^6 m^3
- DEPTH TO WATER TABLE IN TAILINGS - 1 m

around 0.1% uranium. The particle size distribution of the tailings is such that approximately 50% of the tailings solids would be less than 75 um in diameter. Tailings are assumed to contain 4 to 5% of the original uranium (i.e. between 95 and 96% of the uranium was extracted from the tailings during the milling process). Essentially all of the original radium in the ore typically remains with the tailings material. The radium concentration was selected to range between 5.0 and 15 Bq/g. Of this amount, approximately 25 to 50% of the radium content is associated with gypsum which forms as a result of adding lime (or limestone) to the tailings to neutralize the residual acid carried over from the leaching circuit. The remaining radium content is associated primarily with the fine grained tailings particles. The amount of gypsum (as $CaSO_4$) in the tailings was assumed to be 5% of the tailings mass.

All surface runoff from the tailings area drainage basin flows across the tailings surface thus contacting the tailings solids. A schematic representation of the flow distribution within the tailings area is presented on Figure 2.2. The tailings drain towards a conventional earth-filled water-retaining structure, at the outlet of the valley. Excess runoff from this area is decanted via a spillway to a small creek which flows into a river 200 m from the dam. The surface runoff component was assumed to equal 85% of the net precipitation on the tailings basin. This includes water which is either decanted via the spillway or percolates through the dam. The remaining 15% was considered to enter the tailings porewater, move downward into a groundwater flow system and eventually emerge at the river 200 m from the tailings site. The transport of radionuclides through the geologic material along the groundwater flowpath was assumed to be retarded by adsorption/precipitation processes. The effect of this assumption was to delay the time to emergence of radium in groundwater at the point of mixing with river water.

The surface and groundwater flow components from the tailings site are immediately diluted with freshwater flow in the river from the larger watershed shown on Figure 2.3. Flow in the river enters a lake some 2 km downstream.

FIGURE 2·2

CROSS-SECTIONAL VIEW OF REFERENCE TAILINGS SITE
SHOWING THE GROUNDWATER AND SURFACE WATER FLOW COMPONENT

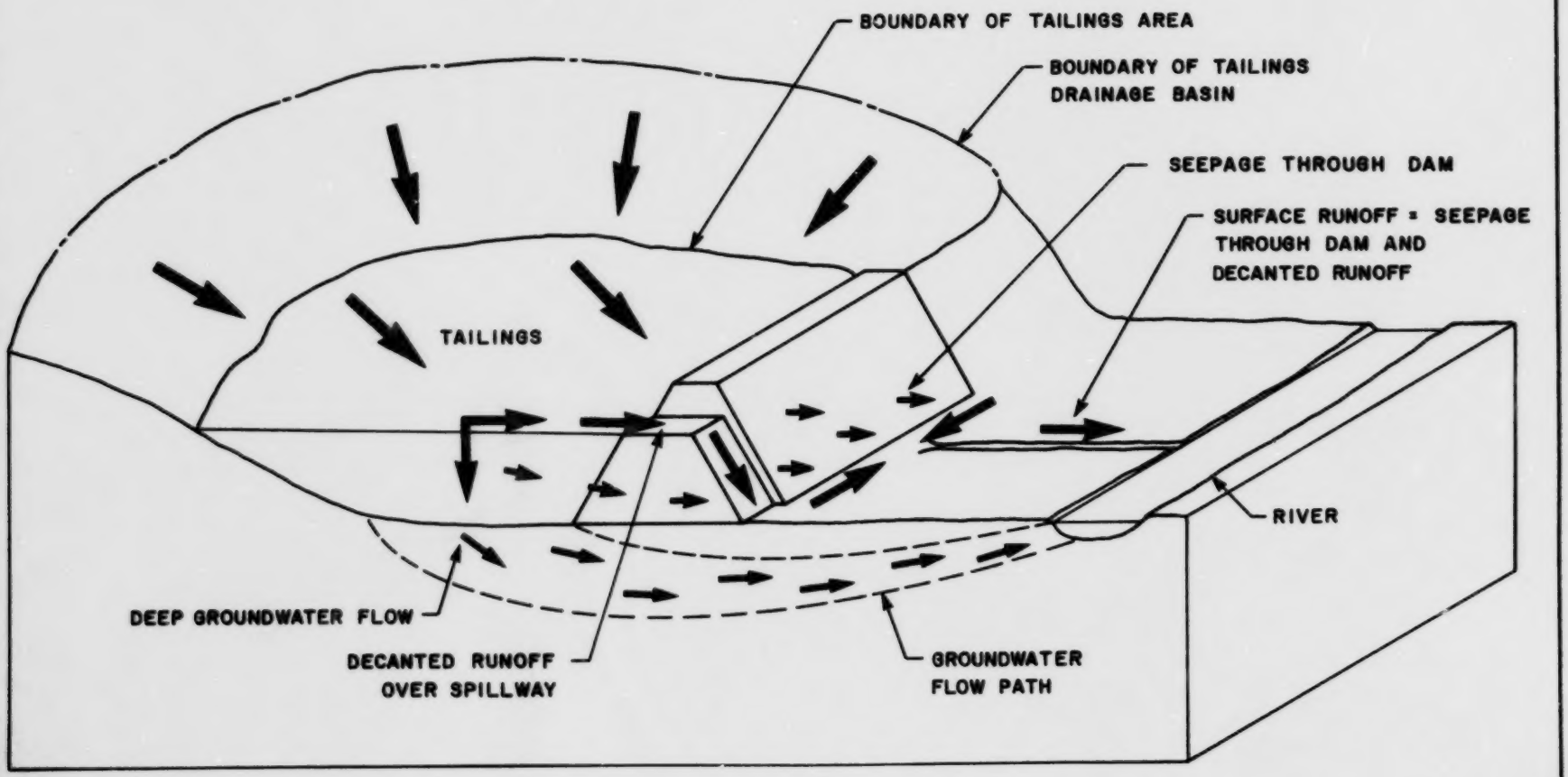
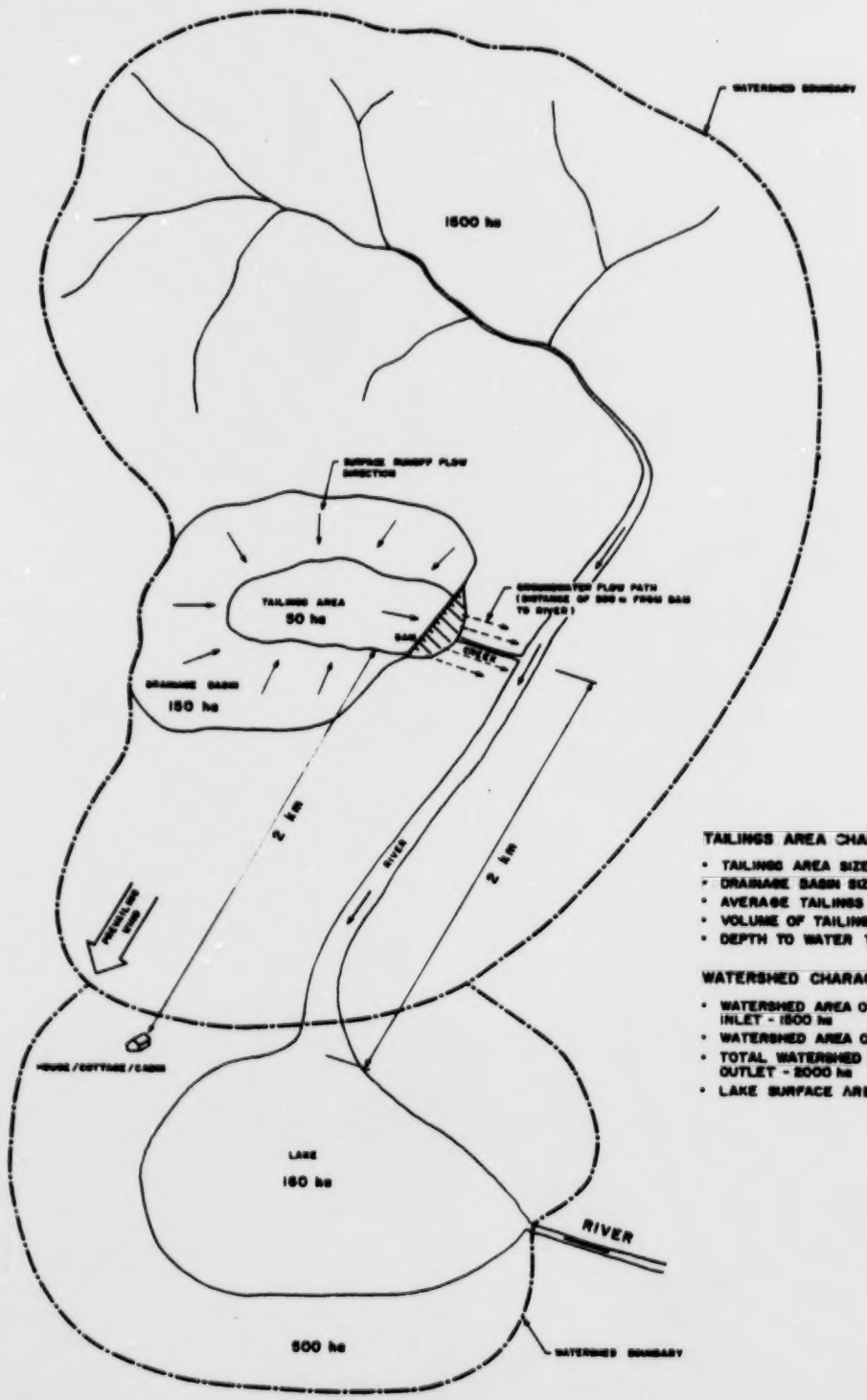


FIGURE 2-3

SCHEMATIC ILLUSTRATION OF WATERSHED SETTING FOR THE REFERENCE TAILINGS SITE



TAILINGS AREA CHARACTERISTICS:

- TAILINGS AREA SIZE - 50 ha
- DRAINAGE BASIN SIZE - 150 ha
- AVERAGE TAILINGS DEPTH - 10 m
- VOLUME OF TAILINGS - 5×10^6 m^3
- DEPTH TO WATER TABLE IN TAILINGS - 1 m

WATERSHED CHARACTERISTICS:

- WATERSHED AREA OF RIVER TO LAKE INLET - 1500 ha
- WATERSHED AREA OF LAKE - 500 ha
- TOTAL WATERSHED AREA TO LAKE OUTLET - 2000 ha
- LAKE SURFACE AREA - 160 ha

The prevailing wind direction follows the direction of the river flow so that the area of maximum potential exposure is both downwind and downstream of the tailings site. The tailings area was assumed to be exposed to normal wind erosion forces.

2.3 Environmental Pathways of Exposure to Man

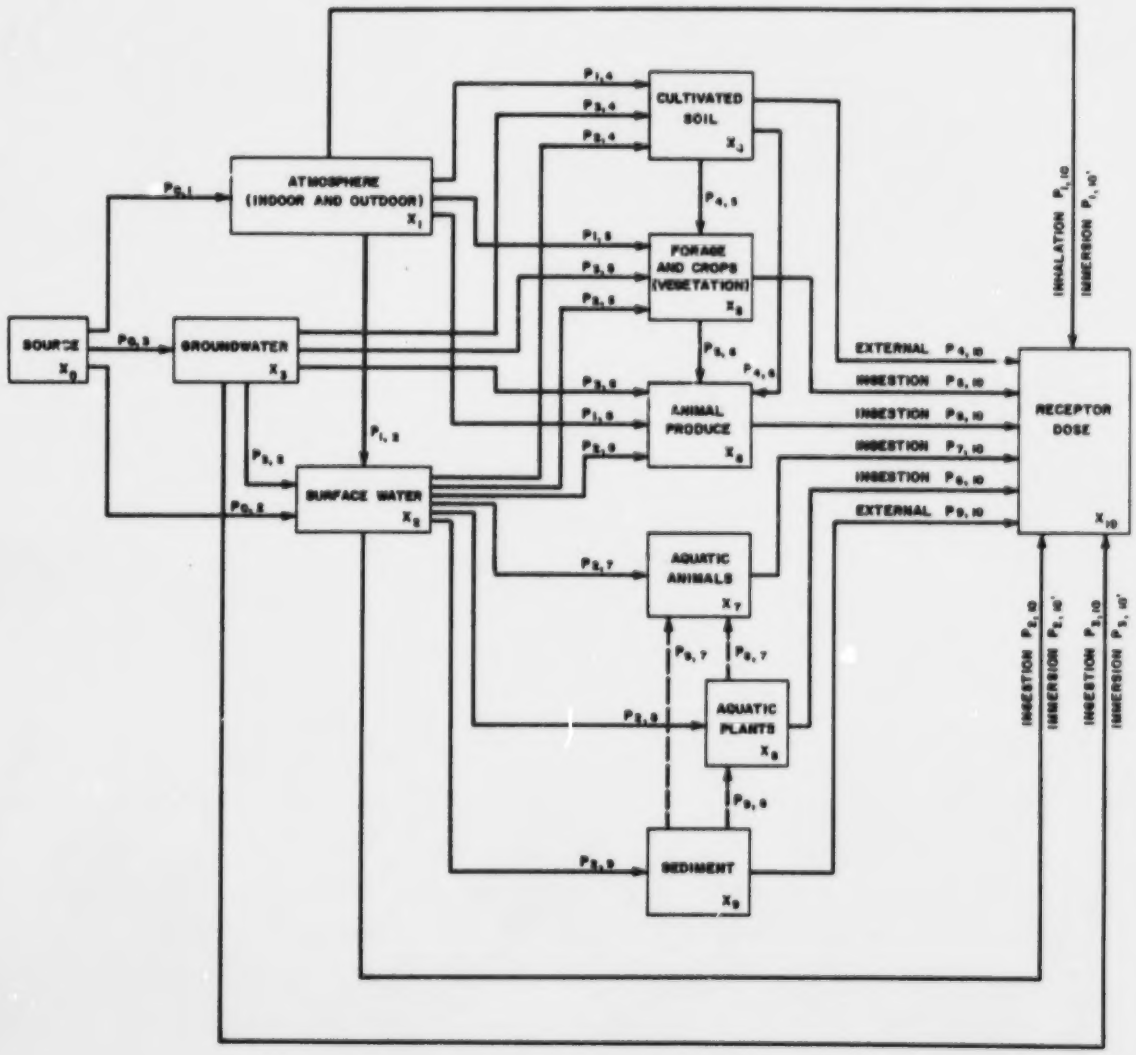
Pathways analysis provides the link between the release of radioactive or non-radioactive contaminants to the environment and the exposure of man to these contaminants. In pathways analysis, the passage of a contaminant through the environment is traced along the following sequence of steps: release to the environment, transport through the air and water, uptake by biota, and human exposure. Source characterization, local atmospheric and hydrologic conditions, radionuclide transfer factors, population distribution and receptor living habits are all important components of pathways analysis.

A generalized pathways analysis diagram is presented in Figure 2.4. The diagram shows several pathways (solid lines) which can be modelled for any given situation. The dotted lines indicate pathways that are implicitly, rather than explicitly modelled. Selection of appropriate pathways depends on various factors including the characteristics of the contaminants, the contaminant source, the environment and the receptor(s) under consideration.

The concept of the analysis shown on Figure 2.4 is that the dose to a receptor for any pathway is calculated by a series of multiplicative steps. Each individual exposure pathway can be represented as a multiplicative chain. The dose received from several pathways is determined by summing the dose estimates from individual pathways. The blocks or compartments on the figure represent either the contaminant emission rate (i.e. the source term - X_o block), the environmental concentration (e.g. the atmospheric, surface water, forage and crops, etc. - X_1 to X_9 compartments) or the dose estimate (i.e. receptor dose - X_{10} block). The solid or dashed lines connecting the various compartments represent transfer models (i.e. lines $P_{0,1}$ through $P_{9,10}$) which take into account such factors as dilution or dispersion, adsorption, decay or

FIGURE 2·4

GENERALIZED PATHWAYS MODEL



ingrowth, etc. Hence the concentration in any of the environmental compartments or (in the case of X_{10}) the dose to the receptor is calculated as:

$$X_j = \sum_i P_{ij} X_i \quad (i \neq j)$$

For example, the concentration of a contaminant in the soil at the receptor site is given by:

$$X_4 = P_{1,4}X_1 + P_{3,4}X_3 + P_{2,4}X_2$$

The definitions and units assigned to the compartment and transfer models shown on Figure 2.4 are summarized on Tables 2.1 and 2.2 respectively.

To demonstrate the probabilistic assessment technique, the number of exposure pathways selected for detailed analysis was limited to six key pathways. The six chosen were those which had been shown in previous tailings management investigations by the study team to be the most significant pathways of exposure for radon and radium. A condensed pathways diagram showing the six selected pathways is presented on Figure 2.5 and includes:

Source	environmental transfer	dose
1) radon exhalation	to atmospheric dispersion	to receptor inhalation
2) dust erosion	to atmospheric dispersion	to receptor inhalation
3) dust erosion	to vegetation uptake	to receptor ingestion
4) dust erosion	to atmospheric dispersion to vegetation uptake to animal produce	to receptor ingestion
5) radium leaching	to surface water and groundwater	to receptor ingestion
6) radium leaching	to surface water and groundwater to aquatic animals	to receptor ingestion

Table 2.1

Pathways Analysis: Model Compartments and Units

Designation	Description	Units
X_0	Source	$Bq.s^{-1}$
X_1	Atmosphere	$Bq.m^{-3}$
X_2	Surface Water	$Bq.m^{-3}$
X_3	Groundwater	$Bq.m^{-3}$
X_4	Soil	$Bq.m^{-2}$
X_5	Vegetation	$Bq.g^{-1}$
X_6	Animal Produce	$Bq.g^{-1}$
X_7	Aquatic Animals	$Bq.g^{-1}$
X_8	Aquatic Plants	$Bq.g^{-1}$
X_9	Sediment	$Bq.m^{-2}$
X_{10}	Dose	$Sv.a^{-1}$

Table 2.2

Pathways Analysis: Transfer Parameters and Units

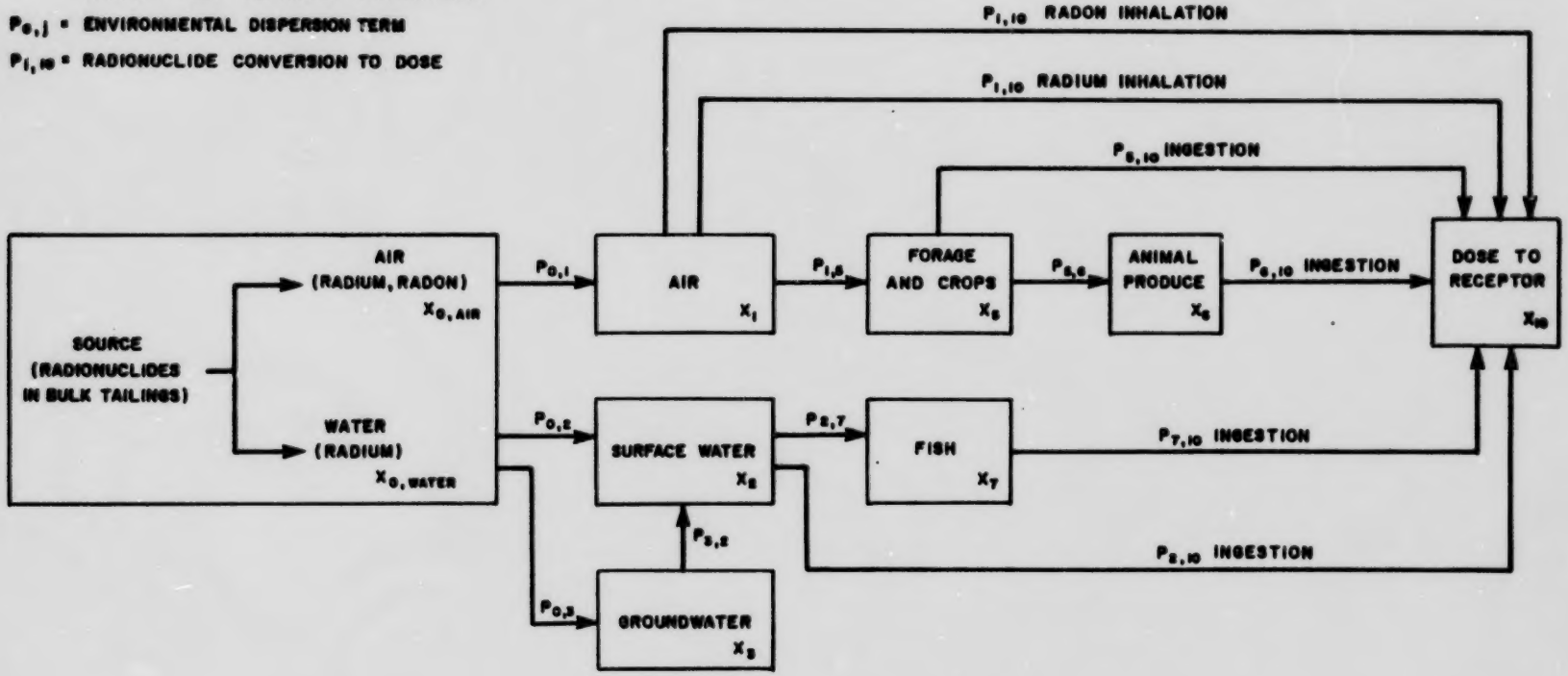
Designation	Description	Units
$P_{0,1}$	Source to Atmosphere	$s.m^{-3}$
$P_{0,2}$	Source to Surface Water	$s.m^{-3}$
$P_{0,3}$	Source to Groundwater	$s.m^{-3}$
$P_{1,4}$	Atmosphere to Soil	$m^3.m^{-2}$
$P_{1,5}$	Atmosphere to Vegetation	$m^3.g^{-1}$
$P_{1,6}$	Atmosphere to Animal Produce	$m^3.g^{-1}$
$P_{1,10}$	Atmosphere to Dose (Inhalation)	$Sv.a^{-1}.Bq^{-1}.m^3$
$P_{1,10'}$	Atmosphere to Dose (Immersion)	$Sv.a^{-1}.Bq^{-1}.m^3$
$P_{2,4}$	Surface Water to Soil	$m^3.m^{-2}$
$P_{2,5}$	Surface Water to Vegetation	$m^3.g^{-1}$
$P_{2,6}$	Surface Water to Animal Produce	$m^3.g^{-1}$
$P_{2,7}$	Surface Water to Aquatic Animals	$m^3.g^{-1}$
$P_{2,8}$	Surface Water to Aquatic Plant	$m^3.g^{-1}$
$P_{2,9}$	Surface Water to Sediment	$m^3.m^{-2}$
$P_{2,10}$	Surface Water to Dose (Ingestion)	$Sv.a^{-1}.Bq^{-1}.m^3$
$P_{2,10'}$	Surface Water to Dose (Immersion)	$Sv.a^{-1}.Bq^{-1}.m^3$
$P_{3,2}$	Groundwater to Surface Water	$m^3.m^{-3}$
$P_{3,4}$	Groundwater to Soil	$m^3.m^{-2}$
$P_{3,5}$	Groundwater to Vegetation	$m^3.g^{-1}$
$P_{3,6}$	Groundwater to Animal Produce	$m^3.g^{-1}$
$P_{3,10}$	Groundwater to Dose (Ingestion)	$Sv.a^{-1}Bq^{-1}.m^3$
$P_{3,10'}$	Groundwater to Dose (Immersion)	$Sv.a^{-1}Bq^{-1}.m^3$
$P_{4,5}$	Soil to Vegetation	$m^2.g^{-1}$
$P_{4,6}$	Soil to Animal Produce	$m^2.g^{-1}$
$P_{4,10}$	Soil to Dose (Groundshine)	$Sv.a^{-1}.Bq^{-1}.m^2$
$P_{5,6}$	Vegetation to Animal Produce	$g.g^{-1}$
$P_{5,10}$	Vegetation to Dose	$Sv.a^{-1}.Bq^{-1}.g$
$P_{6,10}$	Animal Produce to Dose	$Sv.a^{-1}.Bq^{-1}.g$
$P_{7,10}$	Aquatic Animal to Dose	$Sv.a^{-1}.Bq^{-1}.g$
$P_{8,10}$	Aquatic Plant to Dose	$Sv.a^{-1}.Bq^{-1}.g$
$P_{9,10}$	Sediment to Dose (Groundshine)	$Sv.a^{-1}.Bq^{-1}.m^2$

FIGURE 2·5

SELECTED PATHWAYS FOR PHASE I INVESTIGATIONS

NOMENCLATURE:

- X_0 = APPROPRIATE SOURCE RELEASE RATE
- X_j = CONCENTRATION OF RADIONUCLIDE IN ANY ENVIRONMENTAL COMPARTMENT $j = 1, 7$
- $P_{i,j}$ = TRANSFER OF RADIONUCLIDE FROM COMPARTMENT i TO j WHERE i = ANY PREVIOUS COMPARTMENT
- $P_{0,j}$ = ENVIRONMENTAL DISPERSION TERM
- $P_{i,10}$ = RADIONUCLIDE CONVERSION TO DOSE



Within the time frame of the analysis (i.e. the initial 200 year period), the contribution of radium transport through the groundwater pathway is insignificant and was set equal to zero in the probabilistic assessment of individual pathways. The groundwater component was analyzed independently of the pathways analyses however, to demonstrate the effects of time dependency on radionuclide transport (see Section 4.4).

2.4 Reference Receptor

The reference receptor is assumed to be an adult male (Reference Man as defined in ICRP 23, 1975) who lives year-round in a cabin on the lake, as shown on Figure 2.3, some 2 km downstream of the reference tailings site. This individual is assumed to live off the land deriving approximately 50 percent of his food supply from local sources: eating some vegetables and berries grown or harvested near the residence, eating some fish taken from the lake, and eating some caribou hunted in the immediate area. All drinking water is taken from the lake and consumed directly without special treatment for radium removal.

3.0 SYSTEMS CODE AND COMPONENT MODELS

3.1 Introduction

3.1.1 Historical Perspectives

As mentioned previously, mathematical models have been widely used to describe and simulate the environmental behaviour of radionuclides. In fact, compliance with radiation protection standards is often determined by models which convert measurable physical quantities (such as radionuclide concentrations, and atmospheric or geological descriptors) into an estimate of the dose-equivalent to some receptor group. Traditionally, these models have combined single estimates of a number of parameters to arrive at a single estimate of some environmental consequence. Generally, there is some lack of certainty about the 'best' estimate to use for most parameters, and conservative or 'worst case' estimates have often been employed. The deterministic result thus produced is often reported as a conservative estimate of some consequence, with little consideration given to the uncertainty or excessive conservatism associated with it.

In the late 1960's, partly motivated by predictive model results which exceeded certain regulatory guidelines, Fultyn (1968) began examining ways to quantify the uncertainties associated with environmental model predictions. Because of the practical difficulties and expense involved in constantly comparing model predictions with experimental data, other methods of predicting uncertainty were examined. The work of Fultyn (1968), Stocum (1968) and O'Neill (1971), represent early applications of Monte Carlo methods or probabilistic analysis to the environmental behaviour of radionuclides. These approaches combined the uncertainties associated with model parameters to analytically or numerically produce a distribution of expected results rather than a single value.

While this probabilistic approach was new to radioecology, it had previously been used extensively in other branches of science. The Monte Carlo method,

as defined by Meyer in 1956, involved the use of sampling devices based on probabilities, to approximate the solution of mathematical or physical problems. Even at that time it was not a new method. Proceedings of an IBM Computation Seminar in November 1949 noted that the Monte Carlo method had been used by statisticians for over 50 years (Meyer, 1956). One of the early major practical applications of the Monte Carlo method was the design of atomic weapons at Los Alamos during the war. Other applications included problems in gamma ray diffusion, matrix inversion, numerical integration and the solution of differential equations.

As early as 1959, an entire book was written on probability and statistics for business decisions (Schlaifer, 1959) but specific applications were slow to follow. Although the techniques were not new, their application in fields other than physical science really only developed with the proliferation of inexpensive, high-speed computers. Besides the radioecological applications mentioned earlier, probabilistic analysis has begun to find numerous applications in science, engineering, business and social science. Reilly (1969) has applied an opinion analysis program (OPANAL) to both business decision making and scientific problems. Probabilistic assessment has been widely used in the evaluation of different weapons systems and other ordinance problems (Wimmenauer, 1971). It also has found applications in reactor safety studies. For example, SAMPLE and STADIC are computer programs which were used to examine the probabilities of nuclear power plant accidents (Cairns and Fleming, 1977).

Computer programs for the probabilistic assessment of various environmental problems are becoming common. As an illustration of this observation it can be noted that Hoffman and Gardner have explored a wide variety of environmental models using various probabilistic assessment codes developed at Oak Ridge National Laboratories (ORNL) (e.g. PRISMC (Gardner, 1984)). Another major American research effort is underway at Carnegie Mellon University (CMU). The efforts at CMU have included the development of DEMOS, a user-friendly, interactive computer system for probabilistic analysis (Henrion, 1981). Even more relevant to the present application, probabilistic

techniques have been used to analyze the slope stability of some tailings embankment dams in the U.S. (Utah State University, 1982).

The Canadian nuclear fuel waste management program (AECL) is also making extensive use of the technique in their assessment of nuclear fuel waste disposal methods (Dormuth and Sherman, 1981). The work of AECL has focussed on the development of SYVAC, a complex computer code capable of a variety of applications. It has been applied to the study of sub-seabed disposal of nuclear waste as well as various methods of immobilizing nuclear waste (Guvanasen et al, 1983). The Atomic Energy Control Board has also employed probabilistic techniques to examine the possible environmental effects of H_2S releases from a heavy water plant in Ontario (Monserco, 1983).

More recently, the American Meteorological Society recommended that a probabilistic approach be used to improve air quality modelling predictions (Fox, 1982). On a completely different note, Monte Carlo methods were recently advocated for use in Business Venture (risk) analysis (Macaluso, 1984).

The previous examples were cited to illustrate that probabilistic assessment is far from a new technique. It has been successfully applied to a variety of models in many fields. There are no obvious reasons why the technique should not be equally applicable to the study of the long-term environmental effects of uranium tailings areas.

3.1.2 Selection of Appropriate Systems Code

As indicated in the previous section, a number of computer programs are currently available to carry out probabilistic assessments. Since one objective of this study was to demonstrate the probabilistic assessment concept using 'off-the-shelf' technology, several of the available tools were considered for possible implementation. A number of the available programs are quite large and complex and thus are expensive and potentially difficult to implement. These are described briefly below:

PRISMC - developed for the Swedish high level waste program by ORNL and implemented (Gardner, 1984) to evaluate a variety of environmental models. PRISMC employs a Latin Hypercube Sampling strategy (see Appendix E.1 for a discussion of sampling strategies) and can sample from a variety of input distributions. As presently formulated, it requires the use of the IMSL subroutine library and has a number of interesting statistical features.

SYVAC - developed by AECL (eg. Dormuth and Sherman, 1981) for the nuclear fuel waste management program. Although little detailed documentation has been published, SYVAC is known to be a relatively large and complex code with a high degree of internal checking. It has been applied to the prediction of the very long term environmental effects of high level waste disposal technologies.

DEMOS - is the interactive computer system developed at Carnegie Mellon University (Henrion, 1981). It can incorporate a number of different input parameter distributions and is reported to be well integrated with graphics terminals. This allows the interactive generation, display and manipulation of graphical displays of results during a simulation.

The OPANAL program developed by P. Reilly at the University of Waterloo (Reilly, 1969) is relatively simple, and easy to implement. It has been applied successfully to a wide variety of problems in the past 15 years, and is thus well tested and well documented. Since it was also readily available, it was selected for implementation in this study.

OPANAL - is written as a computer subroutine (in FORTRAN IV) with variable dimensions so that it can accommodate any number of input and output parameters. It can sample randomly from six parametric distributions (normal, lognormal, uniform, log-uniform, triangular and beta) and can also sample from any arbitrary distribution

specified as a set of x and y coordinates by the user. The user must also supply the model to be evaluated (in appropriate FORTRAN Code) and an appropriate data file. The program does all the sampling and bookkeeping and produces a set of values for the calculated parameters. It stores all of the parameter values for further statistical analysis and produces simple frequency histograms for all desired parameters (see also Appendix B).

The environmental models used, the Quality Assurance tests performed and the output results produced by OPANAL are described in subsequent sections.

3.2 Environmental Pathways Descriptions

The six environmental pathways considered in this study were briefly outlined earlier. They have been selected partly to provide a reasonable generic estimate of the environmental effects of uranium tailings areas, but, more importantly, to illustrate the application of the probabilistic assessment concept to this problem. The pathways have been simplified for clarity in illustrating the technique and the reader is cautioned that some second order (not necessarily insignificant) effects may not have been included.

Each pathway consists of a source term, a dispersion term and one or more transfer terms. The source term models represent the release rate of the appropriate radionuclide (only radon and radium were considered in this study) from the tailings material into air or water. The dispersion terms model the dilution (including removal processes) of the radionuclides in air or water and their transport to the specified receptor locations. Transfer models mathematically describe the relationship between radionuclide concentrations in adjacent environmental compartments. For instance, given a concentration of radium in the air or water, a transfer factor allows the calculation of the concentration of radium in plants or fish at that location.

The six pathways can be logically combined into three groups; each group with a common source term. Three pathways contain the radium source to air (i.e.

dust erosion) as their initial term, two pathways depend on radium release to water (leaching), while only a single pathway is modelled with radon release to air (emanation) as its source term. The three groups of pathways considered are shown in Figures 3.1 to 3.3. The pathways are described qualitatively in this section; detailed discussions of the component models can be found in Appendix C. As described earlier and indicated on Figure 3.4, the effects of the six pathways can simply be added to arrive at the total dose to the receptor.

3.2.1 Radium in Air

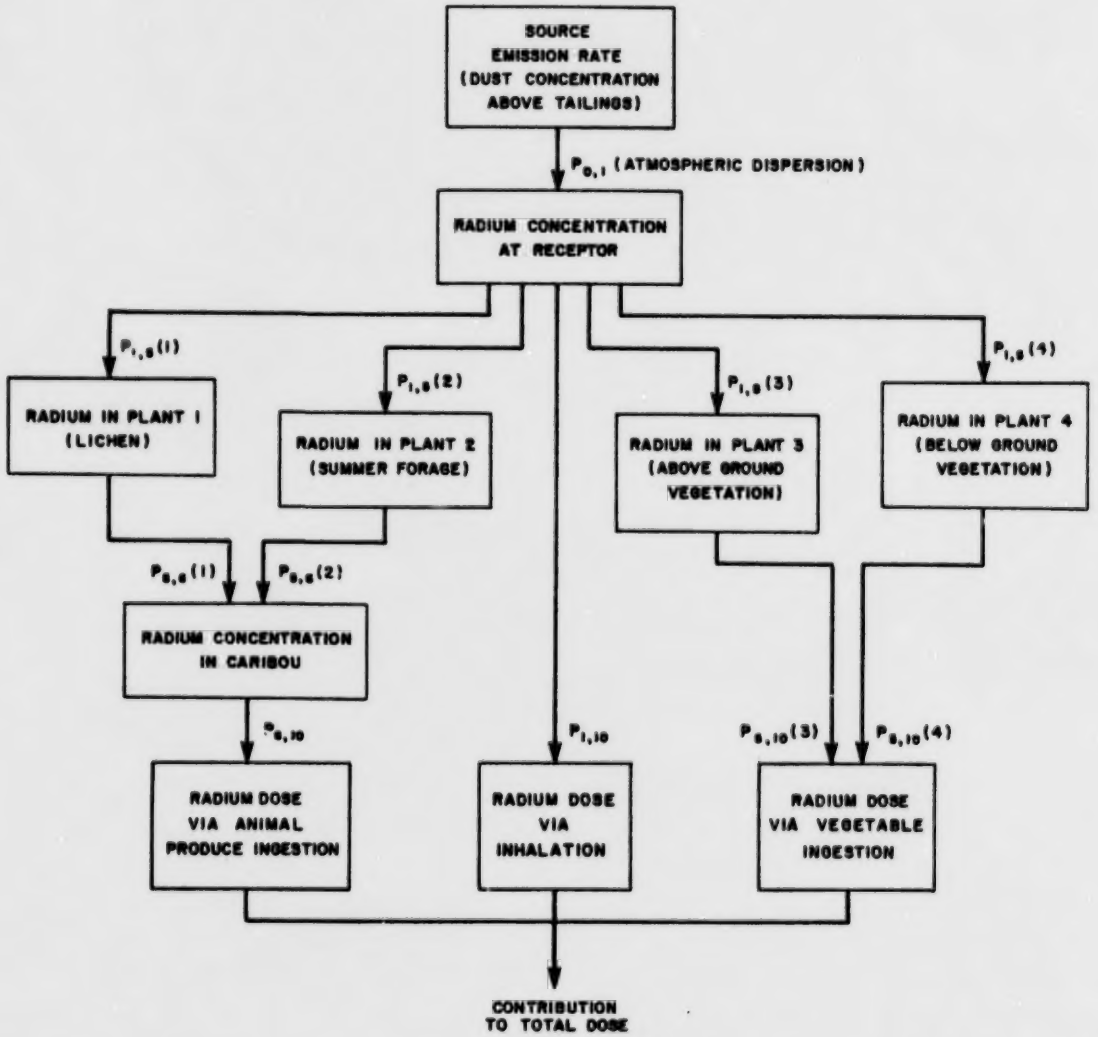
Dust released from the tailings surface as a result of wind action provides the source for the radium dust pathways (Figure 3.1). This source term model is based on a rather poorly defined parameter, the 'resuspension factor'. The subsequent atmospheric dispersion model, which is based on gradient transport theory, is widely used in air dispersion studies. Although the mathematical derivation can become quite complex, the model physically depends primarily on weighted averages of wind speed, turbulence and direction. The distance from the source to the hypothetical receptor site is also specified in this model.

As shown in Figure 3.1, the dust (containing radium) has a variety of fates. First, it may be inhaled directly by the receptor. Alternatively it may settle on a variety of plants assumed to be present at the location. Vegetation considered in this study included lichens, caribou summer forage (a variety of grasses), above ground garden vegetables (eg. lettuce) and below ground garden vegetables (eg. potatoes). Each type of vegetation accumulates radionuclides at a different rate and thus has a unique transfer factor. These vegetables, when consumed by the receptor, will contribute to the total dose.

The receptor, as described earlier, is also assumed to hunt locally for some of his meat. Caribou may be found at some locations and is the only animal considered in this study to be hunted and eaten by the receptor. The caribou

FIGURE 3·1

RADIUM DUST PATHWAYS



NOTE:

P(I) REFERS TO TRANSFER FACTOR P, FOR PLANT TYPE (I) (I = 1,4)

FIGURE 3·2

RADIUM WATER PATHWAYS

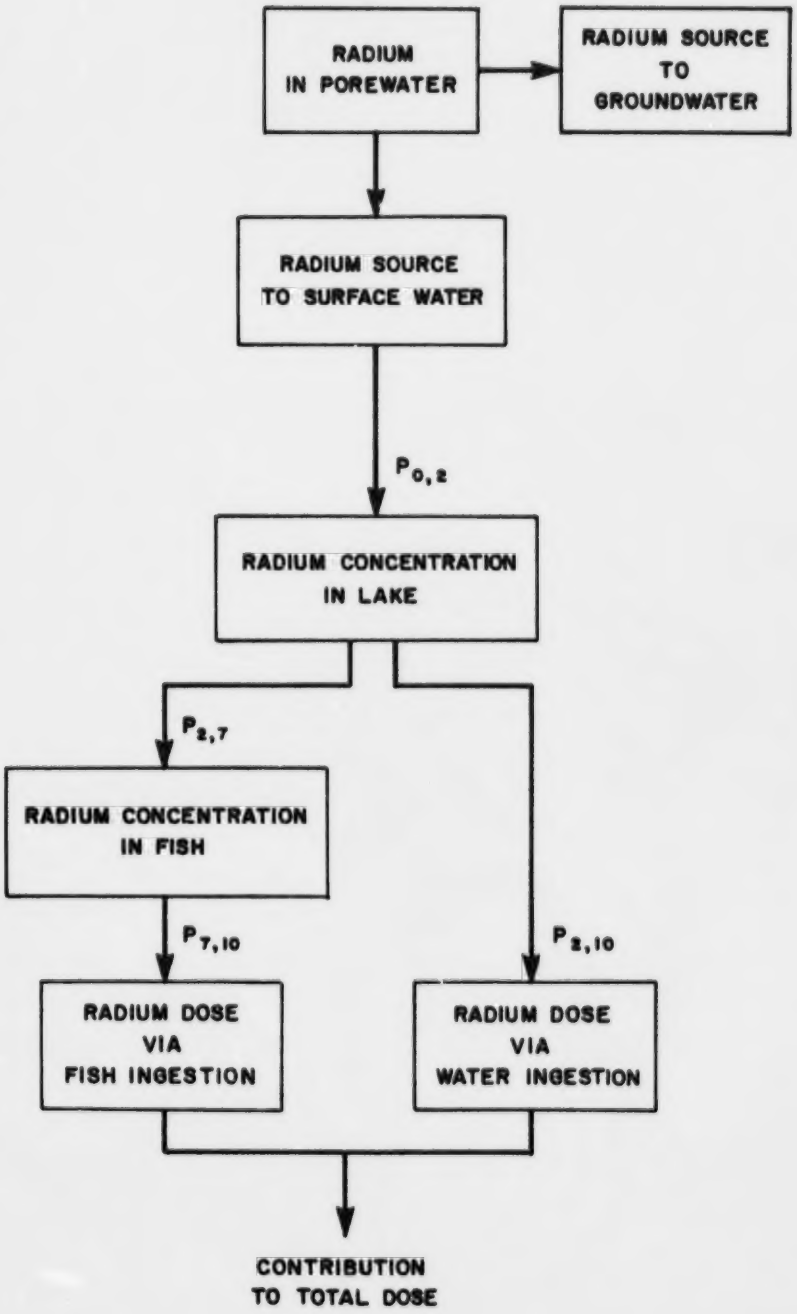


FIGURE 3-3

RADON PATHWAYS

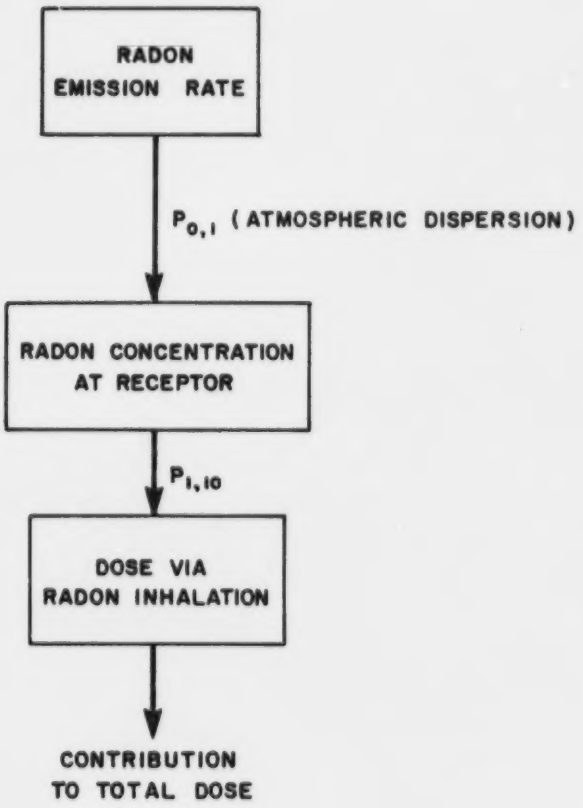
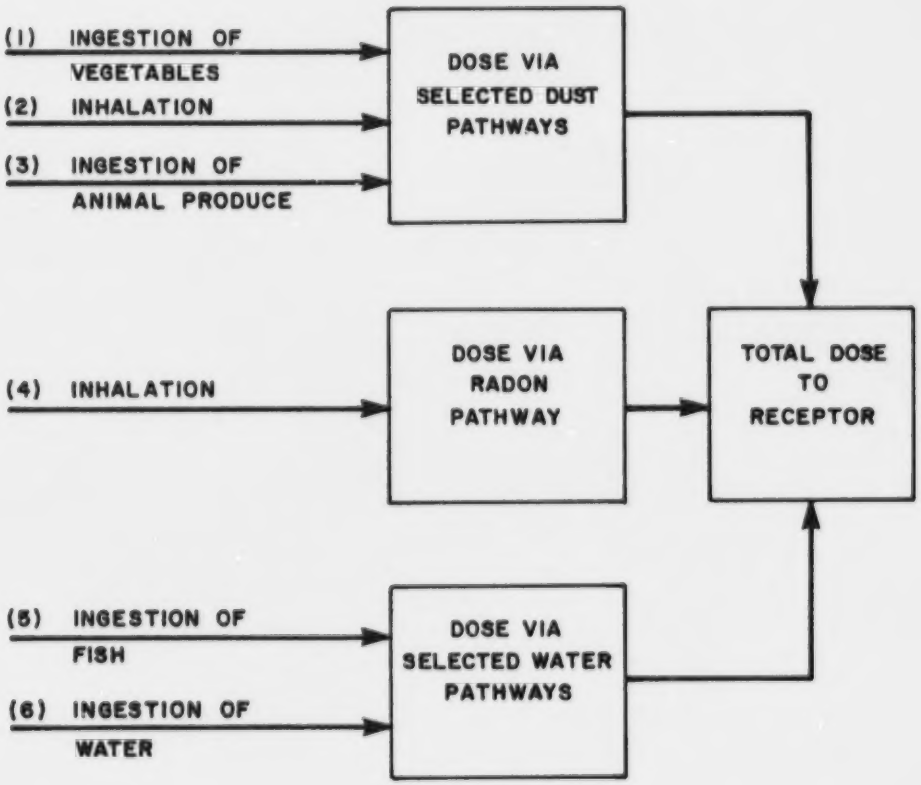


FIGURE 3·4

ADDITIVE CONTRIBUTIONS TO RECEPTOR DOSE



may have consumed some affected vegetation (summer forage or lichens) and thus the caribou meat may contain small amounts of radionuclides. Ingestion of this meat by the receptor will result in a further dose.

3.2.2 Radium in Water

Since radium is initially present in solid form in the tailings, the water source term first models the mobilization of radium from the solid phase. The mobilization essentially involves leaching of the radium from a variety of solid gypsum co-precipitates into the tailings porewater. This leaching rate is a complex function of the solubility of gypsum, the rate of oxidation of the pyritic tailings component and the concentration of sulphates in the tailings porewater. The model predicts a rate of release of radium into the water contacting the tailings mass, either as surface runoff which flows across the tailings surface or as subsurface flow which penetrates the tailings mass, both of which are assumed to take on the characteristics of tailings porewater.

The subsurface flow component moves down to the tailings water table located approximately 1 m below the surface. Some of this subsurface flow may seep out through the dam, while the remainder enters the deeper groundwater regime. The groundwater dispersion term is considered separately because of its time dependent behaviour (see Appendix C and Section 4.4).

The surface water dispersion term models the mixing of the tailings basin runoff with fresh water flow in a river. Additional mixing with freshwater occurs as the river flows downstream and into a lake. Radium removal from the lake water by sorption on suspended inorganic solids and biomass, which settle to the lake bottom, is also modelled. Biodegradation of the organic matter results in some radium being released back into the water column. The lake bottom sediment may thus act as a sink or source of radium. The net result is that the model predicts the concentration of dissolved radium in the lake water column.

As shown on Figure 3.2, two pathways of exposure to the receptor have been considered. All of the receptors' drinking water is assumed to come directly from the lake with no treatment for radionuclide removal. In addition, fish from the lake may also be consumed. Since these fish live in water with an elevated radium content, they too will show an elevated radium level and provide a further contribution to the total dose.

3.2.3 Radon Pathway

The radon pathway (shown in Figure 3.3) is relatively straightforward, involving only a source term, an air dispersion term and a single transfer term. Since radon is a gas, the mechanism of its release into the air differs considerably from that of radium in dust. The mathematical model of radon emission incorporates the rate of production of radon (through the radioactive decay of the parent radium), as well as its rate of escape from the solid tailings material (emanation coefficient) and its transport through the tailings mass to the surface. Several terms are also included to consider the effects of winter freezing conditions and of covering the tailings with uncontaminated material.

The atmospheric dispersion model used is identical to that described earlier. That is, the same conditions are assumed to hold for the dispersion of radon gas in air as for radium bearing dust particles. The single transfer factor acts essentially to convert the concentration of radon gas at the receptor site into a dose to the receptor.

3.2.4 Parameter Value Selection

Once the component models have been derived and defined, values must be assigned to all of the parameters in these models. This is not always an easy task since a variety of parameters are defined in the models. Some, such as those describing the dimensions of the reference tailings site, have constant, arbitrary values. Others are physical constants with generally accepted empirical values, such as radionuclide decay constants. However, a number of

parameters are not as easily characterized.

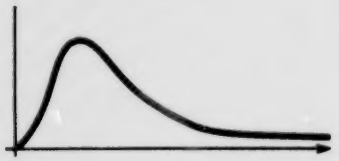
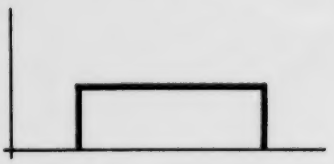
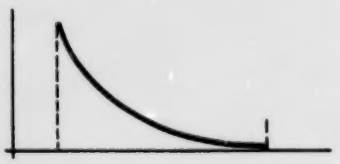
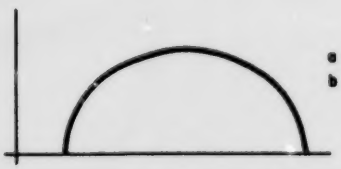
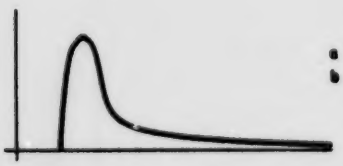
These latter parameters are generally physical quantities, such as annual rainfall, tailings bulk density or wind speed, which have an inherent variability. Often in the past, single, 'conservative' values were assigned to these parameters and a single 'conservative' result obtained. As previously described, the object of a probabilistic evaluation is to assign likely distributions to these parameters, and thus obtain a distribution of likely results. For some parameters this is a relatively straightforward task (eg. the distribution of average annual precipitation which can be obtained from historical records) while for others, the available information may be ambiguous or encompass a wide range of values from various locales. In the latter instance, intuition, physical principles, field sampling, previous experience and professional judgement are used to arrive at appropriate distributions.

Details of the parameter values selected for the Phase I investigation are provided in Appendix D. A number of the distributions selected were based on relatively firm experimental data, while others were selected quite arbitrarily. Since the primary objective of this study was to examine the probabilistic assessment tool, an appropriate estimate was made where information regarding a parameter was not readily available. Also, the distributions were selected to ensure that all of the types of distributions available in the OPANAL program were exercised. In all cases, the distributions selected are believed to be reasonable approximations of reality for the reference site.

The probabilistic assessment tool selected for this study allows six different types of input parameter distributions to be specified. The properties of these distributions are described briefly in Appendix A while the means of specifying them for computer implementation is summarized in Table 3.1. Each of the distributions has its inherent advantages and disadvantages. For example, a number of parameters may be represented by the familiar normal distribution; however, mathematically, this distribution may include negative

TABLE 3·1

AVAILABLE PARAMETER DISTRIBUTIONS IN OPANAL

DISTRIBUTION TYPE	PARAMETERS REQUIRED	TYPICAL SHAPE
• NORMAL	• MEAN • STANDARD DEVIATION	
• LOG NORMAL	• GEOMETRIC MEAN • GEOMETRIC STANDARD DEVIATION	
• UNIFORM	• MAXIMUM • MINIMUM	
• LOG UNIFORM	• MAXIMUM • MINIMUM	
• TRIANGULAR	• MINIMUM • MODE • MAXIMUM	
• BETA	• MINIMUM • MAXIMUM • a, b	MANY POSSIBLE SHAPES
	EXAMPLES:	
		

values, which are often physically impossible. The lognormal distribution, on the other hand, includes only positive values, but also has a very long right hand tail, which occasionally produces unrealistically large values. The beta distribution is particularly interesting, since it can acquire essentially any shape, depending on the values specified for the parameters a & b.

3.3 Quality Assurance: Concepts and Applications

Computer implementation of the probabilistic assessment concept comprises a main random-sampling program and a user-specified model. Since the two routines are very dissimilar, different quality assurance (Q.A.) strategies are required. In the context of computer programs (software), quality assurance has been defined as "A planned and systematic pattern of all actions necessary to provide adequate confidence that the item or product conforms to established technical requirements." (IEEE/ANSI, 1981). This generally requires a systematic approach to both qualitatively examine and quantitatively test the program to ensure that it does what it is expected to do. The quality assurance strategies are considered separately in the following sections.

3.3.1 OPANAL Q.A.

The basic requirement of the probabilistic assessment tool selected for implementation in this study was the capability to randomly pick values of distributed parameters and combine them to produce a range (distribution) of calculated results. For a number of reasons, the OPANAL program developed by Dr. P. Reilly at the University of Waterloo was selected as a likely candidate to demonstrate the probabilistic assessment concept. Since this computer program had been extensively tested and widely used previously, the quality assurance strategy had a rather unique focus, distinctly different from that normally applied to the development of a new software product (as outlined in Appendix E.5 and Section 3.3.2). In this instance, a number of tests were conducted to ensure that the results produced by the program were a function of only the input parameter distributions and the specified mathematical

models and not simply artefacts of the computer routine used.

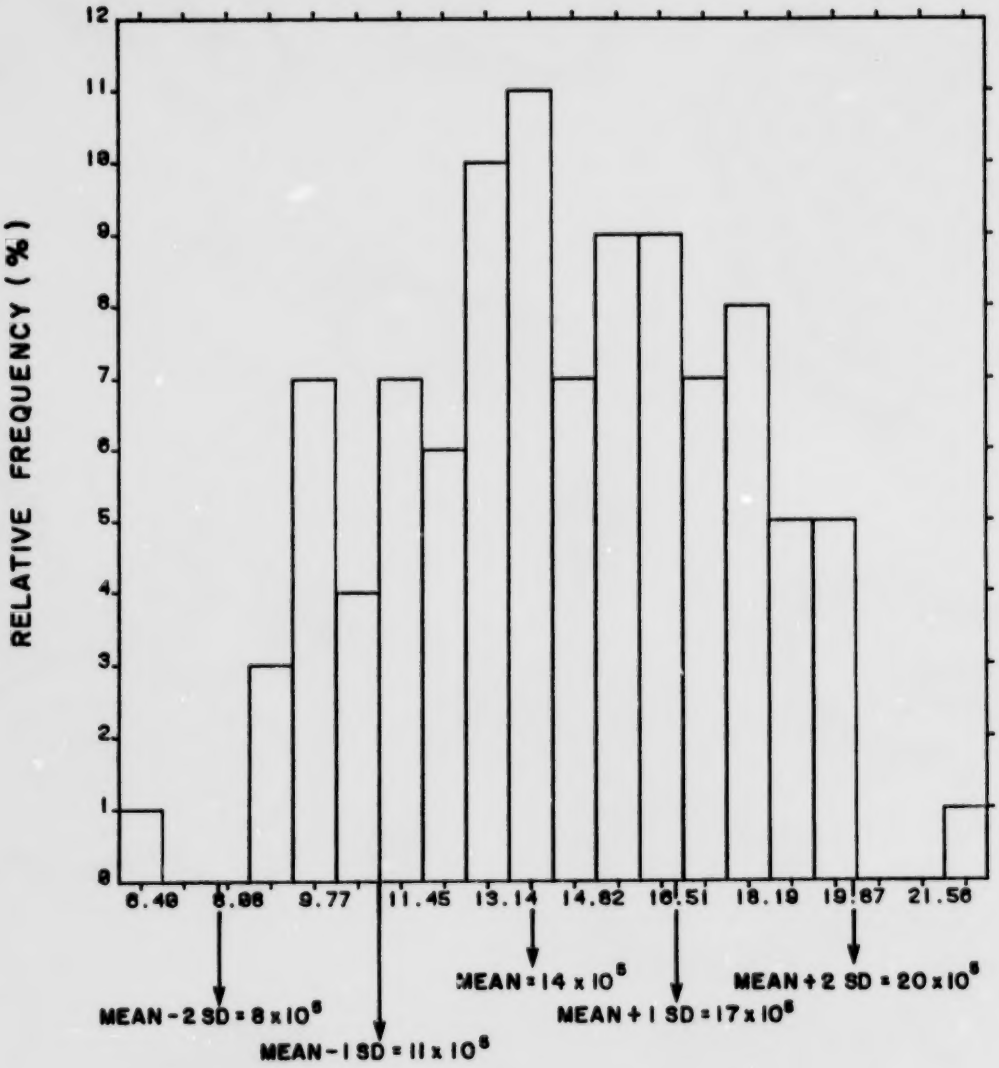
First, the user-specified programs were constructed so that the distributions of sampled values of the input parameters were displayed, along with distributions of significant calculated parameters. Parameters representing each of the six basic types of probability distributions were examined to ensure that the sampling routines were functioning correctly. Representative parameters of each type are displayed in Figures 3.5 to 3.10. The expected shapes of the distributions are presented in Table 3.1, while details regarding the graphical presentation (a frequency distribution) can be found in Appendix A. The results shown were obtained from 100 random samples of each of the distributions. This relatively small number of random trials accounts for the uneven shape of some of the histograms. A greater number of trials would have produced 'smoother' curves. The number of random trials required for various purposes was examined in more detail and is discussed later in this section.

A second attribute of the random sampling routines was also examined. These routines depend on a user-specified 'random number seed' to randomly sample from the specified distributions. Different initial random numbers are expected to produce results that may be noticeably, but not statistically, different. As a test, one model was evaluated using ten different random number seeds. The results were as expected; ten somewhat different sets of values were produced, but these differences were found to be statistically insignificant. That is, the results were essentially random samples from the same underlying population (i.e. variances were homogeneous and means indistinguishable at the 95% level).

As a third test, the OPANAL results were checked to ensure that the program was properly combining the input parameter uncertainties. By considering a purely multiplicative, artificial model of the form: $y = x_1*x_2*x_3*...x_n$, where the means and variances of all the x's were known, an analytical estimate of the mean and variance of y could be made. This was done by first taking logarithms of both sides of the equation, producing an additive model:

FIGURE 3·5

NORMAL DISTRIBUTION



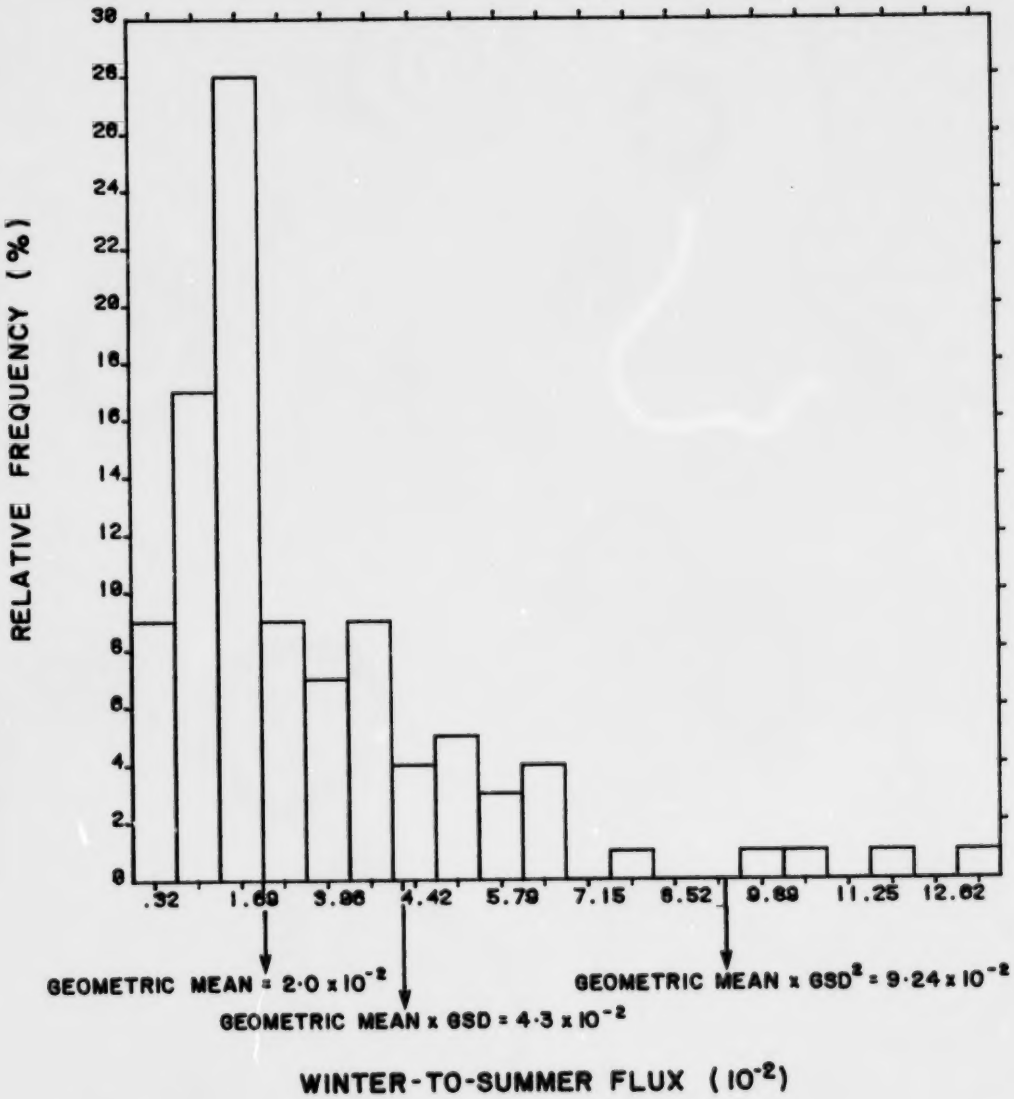
NOTE:

MEAN = 14×10^5

STANDARD DEVIATION (SD) = 3×10^5

FIGURE 3·6

LOG NORMAL DISTRIBUTION



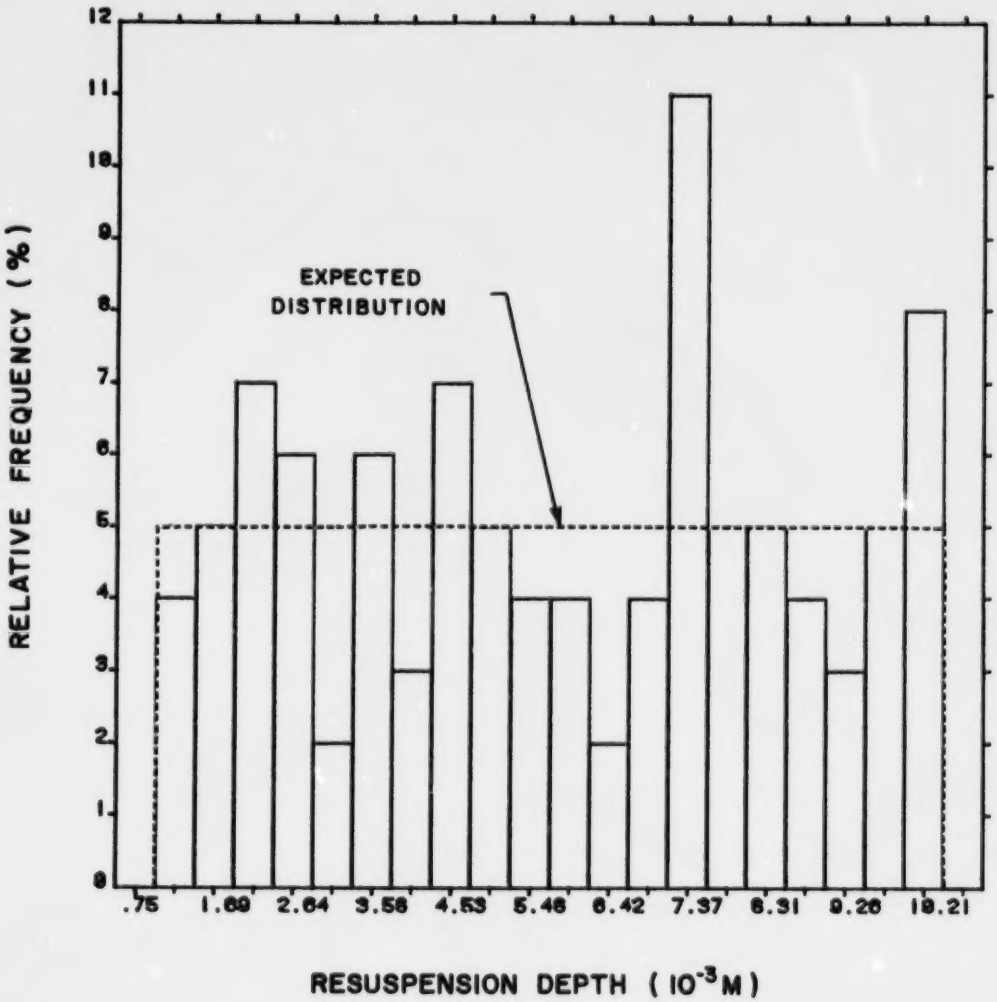
NOTE:

GEOMETRIC MEAN = $2 \cdot 0 \times 10^{-2}$

GEOMETRIC STANDARD DEVIATION (GSD) = 2·15

FIGURE 3·7

UNIFORM DISTRIBUTION



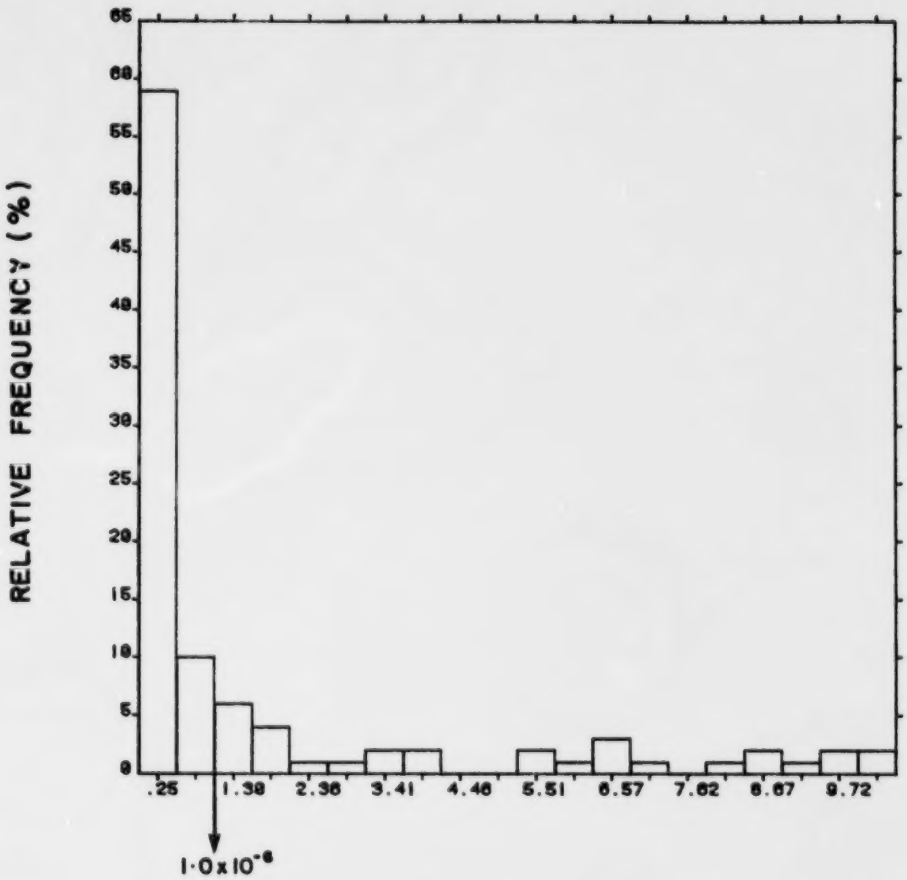
NOTE:

MINIMUM = 1.0×10^{-3}

MAXIMUM = 10×10^{-3}

FIGURE 3·8

LOG UNIFORM DISTRIBUTION



NOTE:

MINIMUM = $1 \cdot 0 \times 10^{-9}$

MAXIMUM = $1 \cdot 0 \times 10^{-5}$

THE LINEAR SCALE DISTORTS THE PRESENTATION OF A LOG UNIFORM DISTRIBUTION. TO EVALUATE THIS FIGURE, CONSIDER FOUR INTERVALS:

i) $1 \times 10^{-9} \leq X < 1 \times 10^{-8}$

ii) $1 \times 10^{-8} \leq X < 1 \times 10^{-7}$

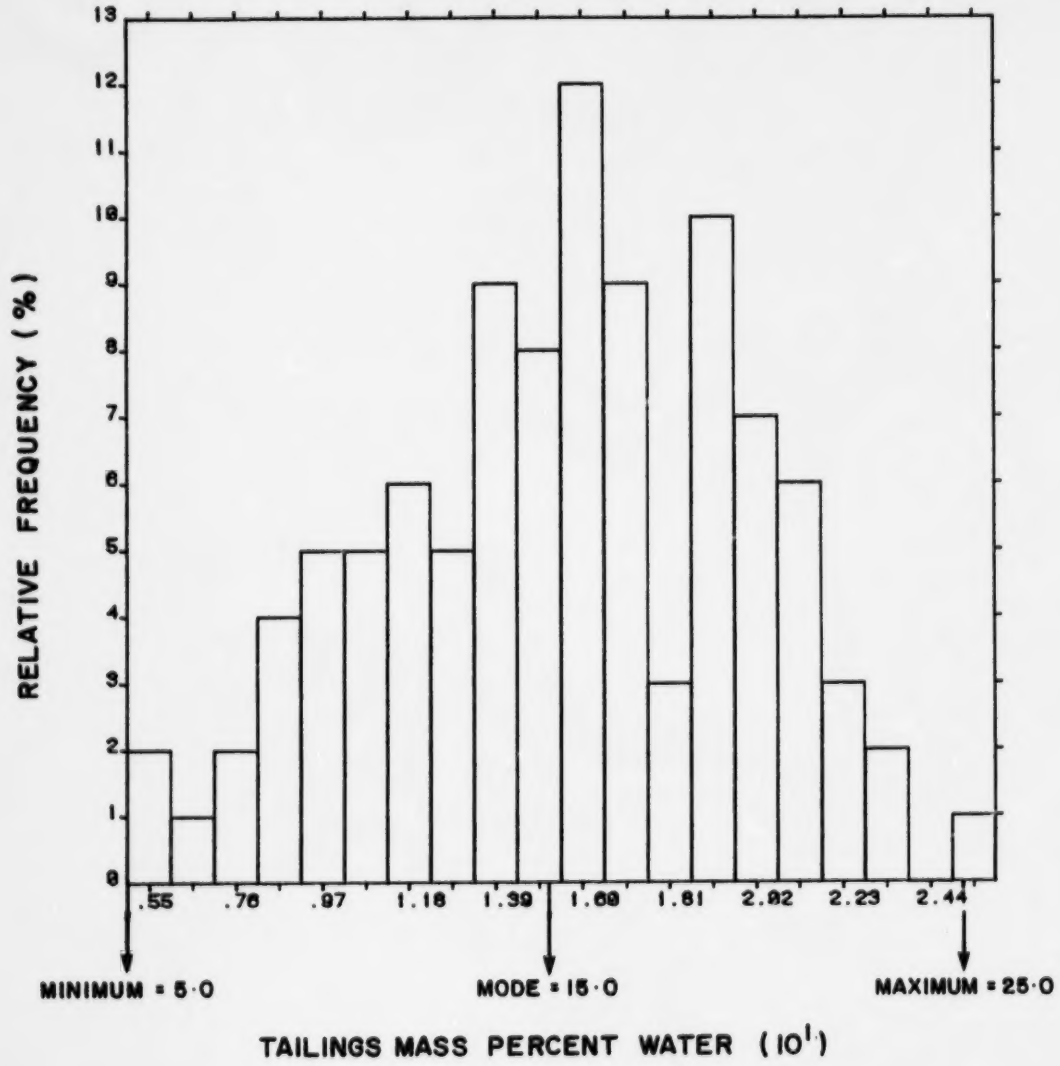
iii) $1 \times 10^{-7} \leq X < 1 \times 10^{-6}$

iv) $1 \times 10^{-6} \leq X \leq 1 \times 10^{-5}$

THE PERCENT OF OCCURRENCES IN EACH INTERVAL SHOULD BE 25. THUS THE PERCENT OF OCCURRENCES LESS THAN $1 \cdot 0 \times 10^{-6}$ SHOULD BE APPROXIMATELY 75. (AS SHOWN).

FIGURE 3·9

TRIANGULAR DISTRIBUTION



NOTE:

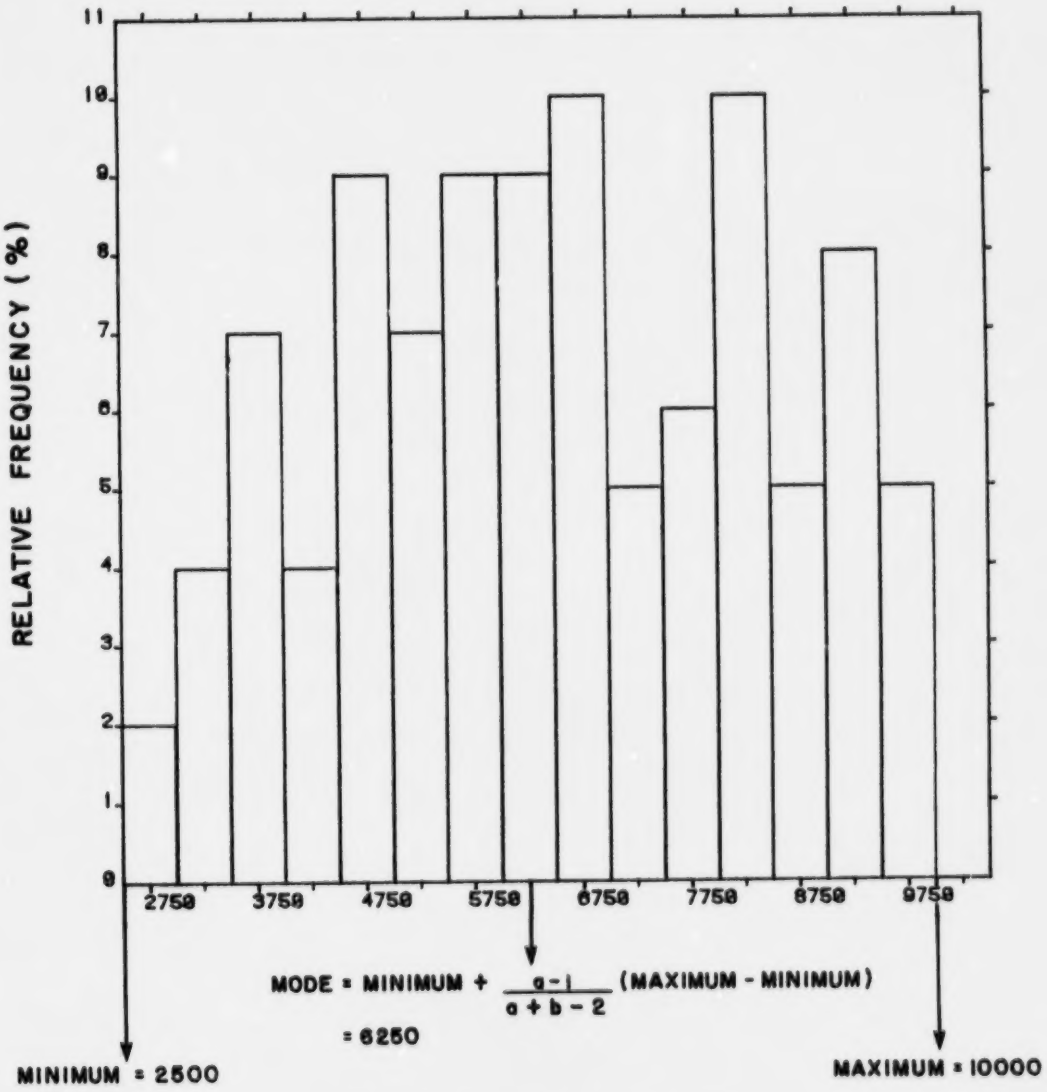
MINIMUM = 5·0

MODE = 15·0

MAXIMUM = 25·0

FIGURE 3·10

BETA DISTRIBUTION



NOTE:

MINIMUM = 2500

MAXIMUM = 10000

a = 1·5

b = 1·5

$$\ln y = \ln x_1 + \ln x_2 + \ln x_3 \ldots + \ln x_n$$

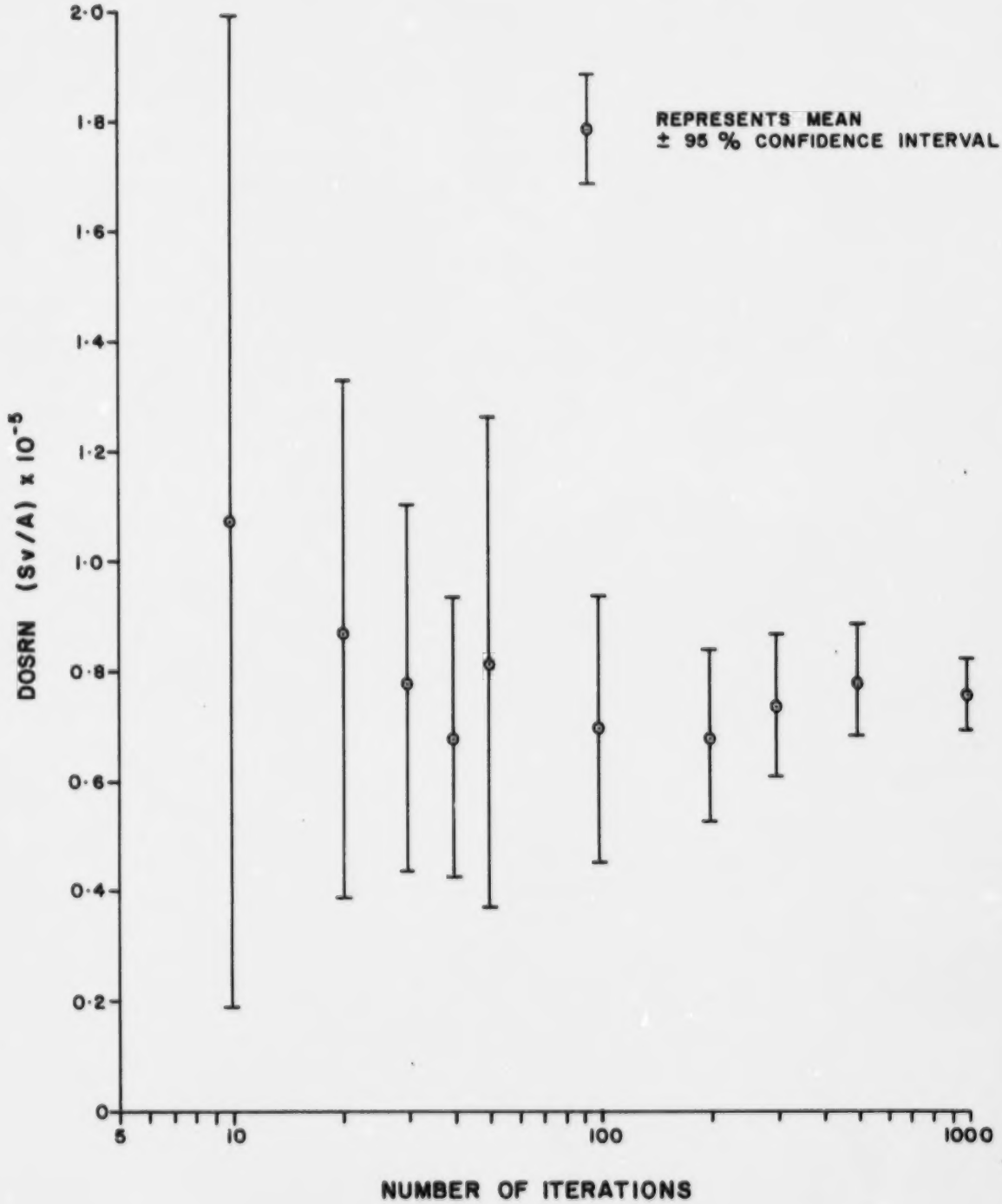
Given this model formulation, the mean and variance of lny are equal to the respective sums of the means and variances of the log transformed values of all the independent parameters. The mathematical estimates of the mean and variance of y were found to correspond very closely to the results produced by an OPANAL evaluation of the same model (i.e. no significant difference at the 95% confidence level). The details of this method were outlined by Hoffman and Gardner (1983).

Finally, the effect of the number of random trials on the output results was examined. These tests attempted to address the questions: "How many trials are enough to estimate: (i) the mean and variance of a calculated parameter? and (ii) the underlying (true) shape of the same calculated parameter?" It was found that 50 to 100 trials were sufficient to produce stable estimates of the mean and variance of a parameter. This is illustrated graphically in Figure 3.11, in which the radon dose to the reference receptor (DOSRN) is plotted against the number of iterations. To produce the results shown, the radon pathways model was evaluated with increasing numbers of trials and the mean and variance of the result were calculated and plotted. As can be seen from the figure, little statistical improvement is gained by performing 1000 trials versus 100. Note also that the increase in statistical accuracy (i.e. decrease in the magnitude of the 95% confidence interval) is similar from 10 to 100 trials and from 100 to 1000 trials (i.e. proportional to $1/\sqrt{n}$ as expected). This indicates that close to 10,000 trials would have to be performed to produce a further increase in accuracy of the same relative magnitude.

It was also found that many more than 100 trials were required to estimate the true shape of an output distribution with some specified degree of confidence. How many runs are actually required for this is still very much a point of debate and depends to a large degree on the desired end-use of the assessment. (See for example Section 4.5 and the topical papers.)

FIGURE 3·11

RADON DOSE MODEL
DOSRN vs. NUMBER OF ITERATIONS



These few tests, as well as the results of several benchmark test cases (supplied by the program developer) provide a reasonable degree of confidence that the OPANAL program is, in fact, performing satisfactorily.

3.3.2 Component Models Q.A.

As outlined earlier, the six pathways modelled were implemented in three separate groups, primarily to facilitate the Q.A. process. These groups of pathways (dust, water and radon) were subsequently combined to produce the 'all-up' or inclusive results. Since the component model, computer code, unlike OPANAL, was developed specifically for this application a more structured approach to Q.A. was used. The approach focussed on error avoidance during all development activities (verification) followed by extensive testing (validation) of the component models.

Each pathway modelled consists of a series of environmental source, dispersion and transfer terms. The derivation/documentation of these individual terms forms a series of mini-specifications (presented in Appendix C). As the first stage in Q.A. these mini-specs were subjected to a series of critical reviews, known as 'walkthroughs'. Walkthroughs involve the program developer and a group of knowledgeable, objective individuals examining a model in a structured, rigorous manner with the objective of discovering errors in the product under review. This technique was found to be very effective and led to the early detection of a wide variety of errors.

Once the mini-specs were reviewed, revised and implemented on the computer system, the Q.A. emphasis shifted to program testing (validation). Adequate program validation generally requires a combination of several test methods, as described in Appendix E.5. Since the component models involved primarily simple, linear logic flows, only 'black box' tests were performed. By definition 'black box' test cases are designed using only the program requirements, with no detailed knowledge of how the requirements are being satisfied by the computer program. It is often useful to design 'black box' test cases that force the program to evaluate extreme ranges of input or

output parameters.

Using these ideas, one set of tests was done by first assigning each input parameter a specific (nominal) mean value, rather than a distribution of values. Then, each parameter was assigned an extreme high and an extreme low value, while all other parameters remained at their nominal values. A number of errors were revealed by checking the results of each of these tests against mathematically expected results, and against results expected by intuition or comparison with appropriate field data. These errors included poorly specified normal distributions (sampling physically unrealistic negative values), poorly specified lognormal distributions (sampling unrealistically large values) and a number of unit conversion errors. Many of the 'black box' test cases were also used in an initial sensitivity analysis (see Section 4.3).

A further set of 'black box' tests involved perturbing the input parameters by small amounts (1%) about their extreme high and low values. These final tests (similar to boundary value analysis described in Appendix E.5) revealed several 'round off' type errors involving small differences between large numbers. Errors of this sort were easily corrected by altering the order of calculations.

Results produced since the completion of testing have revealed no further errors. While this does not prove conclusively that no further errors remain, it does provide a degree of confidence that the programs function as expected over a wide variety of conditions.

4.0 PROBABILISTIC ASSESSMENT RESULTS

4.1 Statistical Concepts for Interpreting Probabilistic Assessment Results

The objective of this study was to demonstrate that a probabilistic analysis is a useful tool for predicting the long term environmental effects of Canadian uranium mill tailings. To demonstrate the concept, a first-cut analysis of a reference tailings site was performed. This section discusses a few of the statistical concepts useful in interpreting the results of such a probabilistic assessment.

Probabilistic assessment results can be regarded as a sample of results drawn randomly from the population of all possible results. These data are commonly presented in the form of a frequency distribution such as that shown in Figure 4.1 a). An arbitrary (but specific) value V_s is also shown. If the frequency curve is normalized such that the total area under it is unity then the area under the frequency curve F(V) to the left of V_s is determined by integrating it from 0 to V_s:

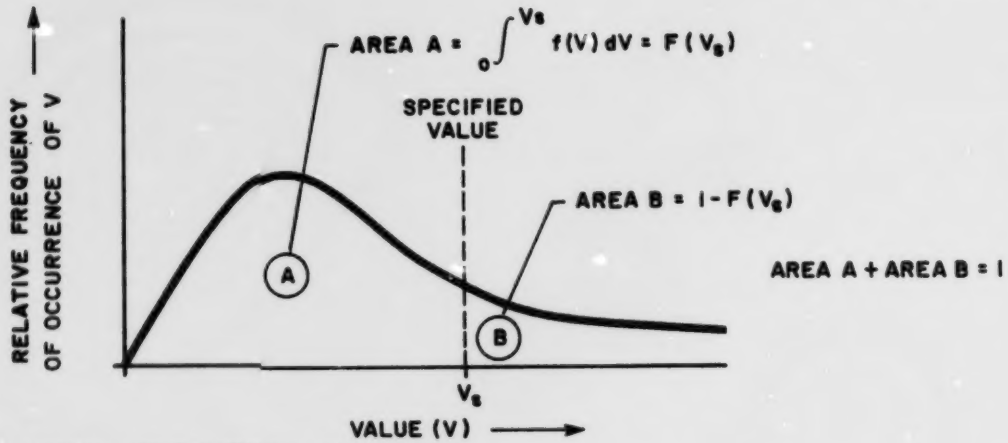
$$\text{area} = F(V_s) = \int_0^{V_s} f(V)\, dV$$

$F(V_s)$ represents the cumulative relative frequency for $V \leq V_s$. The chance (based on observations from the probabilistic assessment) that V is exactly V_s is given by $f(V_s)$.

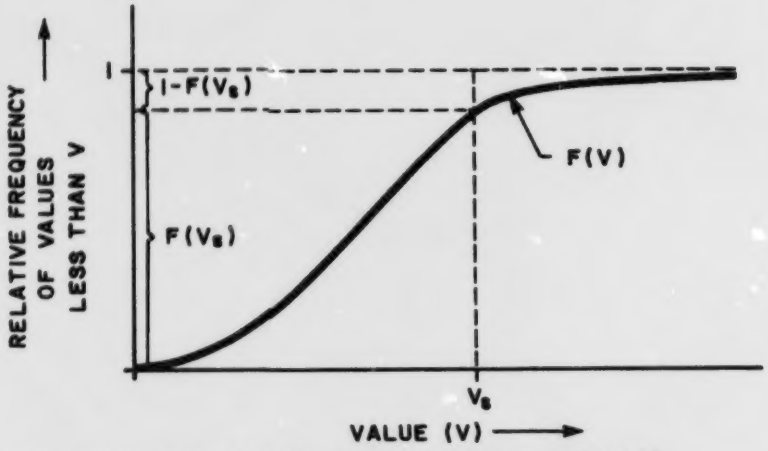
Similarly the chance that $V > V_s$ is given by $1 - F(V_s)$. These concepts are important since current environmental standards tend to be based on specific values (eg. maximum of 5 $mSv.a^{-1}$ whole body dose equivalent to a member of the public) and thus the chance of exceeding a specified numerical standard is an important concept. The more traditional deterministic analysis provides a single result which can be easily compared to the single, hard-and-fast regulatory standard. If the deterministic value falls below the guideline, one can proceed with some confidence. If it is greater than the guideline one

FIGURE 4·1

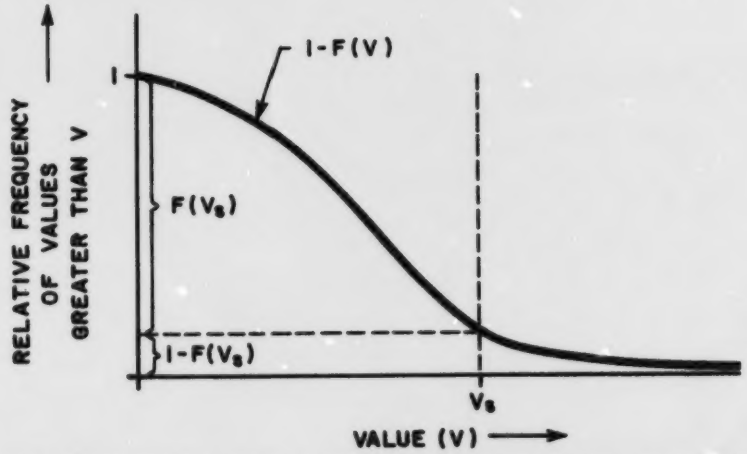
FREQUENCY AND CUMULATIVE FREQUENCY DIAGRAM



A) RELATIVE FREQUENCY DIAGRAM



B) CUMULATIVE RELATIVE FREQUENCY DIAGRAM



C) COMPLEMENTARY CUMULATIVE RELATIVE FREQUENCY DIAGRAM

can either not proceed, adjust the conservatism in the model after due consideration or proceed with some unknown peril. If the probabilistic evaluation indicates that most (e.g. 95%, 99% etc.) of the values can be expected to lie below the limit, then, under most circumstances, it should be possible to proceed. Conversely, if a majority of the values exceed the standard a different response would likely be taken.

Note that the relative frequency histograms already presented (Section 3.3.1) and those that will follow (Section 4.2) are not smooth curves. Because the number of observations is finite, the outline of the frequency diagram is ragged (as a result of sampling error). As the number of observations is increased, the diagrams become smoother and, in the limit, approach the smooth distribution curves. The previous discussion and that which follows is not affected by the curve being approximated by discrete intervals (see also Section 4.5).

Figure 4.1 b) shows a cumulative relative frequency plot which is essentially a plot of F(V) versus V. That is, each point on the graph represents the cumulative relative frequency that the value V will be less than or equal to V_s (i.e. smaller than V_s). It should be noted that the relative frequency curves constructed using present results can be used as estimates of the underlying (population) probability curves.

Figure 4.1 c) extends the concept to the complementary cumulative relative frequency distribution (or downward relative frequency distribution) plot which is determined simply as unity minus the cumulative relative frequency distribution (i.e. 1 - F(V)). A point on this graph represents the cumulative relative frequency that a particular value, V_s, will be exceeded.

The shape of the graphs indicate much about the nature of the distribution. For example a steeply rising cumulative relative frequency distribution indicates that most of the values are clustered around a central point (such as mean or median). A shallow cumulative relative frequency distribution indicates large scatter (dispersion) and hence large uncertainty in the data.

For many purposes, and indeed for almost all applications in this Phase I study, the mean is the most useful descriptor of the outcome of a probabilistic assessment. In addition, it is possible to provide confidence intervals for the mean and also to provide confidence bounds on the entire cumulative relative frequency distribution (or the complementary function). The confidence intervals define an interval within which the value of the estimated, underlying population parameter (mean or cumulative curve) is believed, with a stated percent confidence, to lie. Further discussion of these and other statistical concepts is provided in Appendix A.

4.2 Probabilistic Assessment Results

The probabilistic assessment was performed following the basic outline of the approach discussed in Section 1.2. The assessment was applied to the reference site and selected pathways described in Section 2.0 using the component models and systems code outlined in Section 3.0. This section presents and discusses the results of these calculations. The multiplicative/additive nature of the models used in the study was suited to the calculation of a wide variety of intermediate results during the probabilistic assessment. Several types of output parameter distributions were displayed including the chemical and physical properties of the tailings source, radionuclide concentrations in the various environmental compartments, and factors representing transfer between compartments. The result of primary interest, however, is likely to be the calculated annual radiation dose to the specified receptor.

As discussed in Section 3.2, the total dose can be determined by summation of the contributions from the three pathway groups, that is:

$$\text{total dose} = \begin{pmatrix}\text{dose via}\\ \text{radon}\end{pmatrix} + \begin{pmatrix}\text{dose via}\\ \text{radium in water}\end{pmatrix} + \begin{pmatrix}\text{dose via}\\ \text{radium in air}\end{pmatrix}$$

Since the ultimate effect is the radiation dose to man, this discussion will

focus on four output results; namely, the dose arising from each of the three groups of pathways and the total, combined dose.

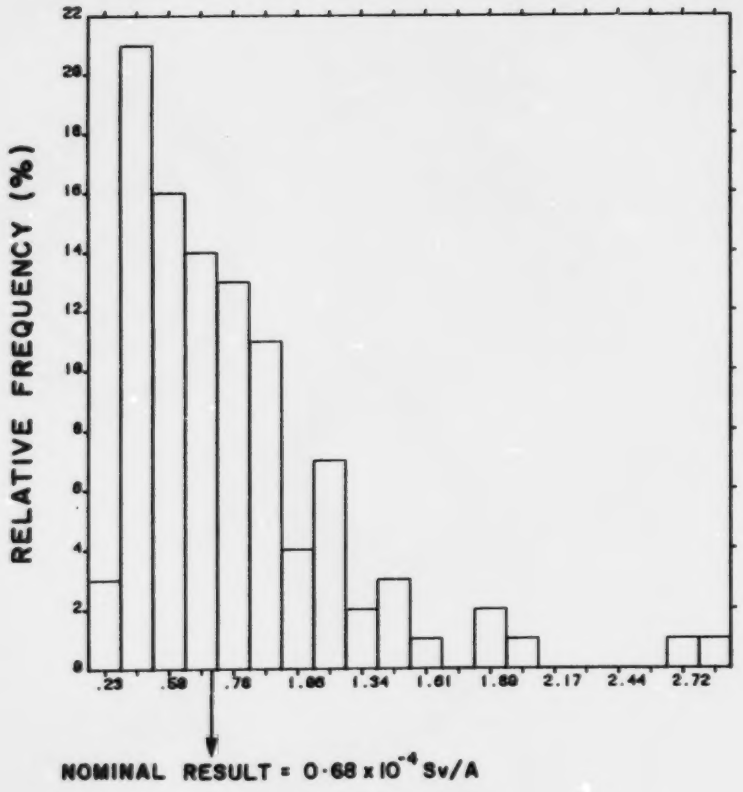
As described in Section 4.1, the results from a probabilistic assessment can be presented in a number of different ways. Since the range of calculated doses may be of interest, as well as some estimate of where the bulk of the calculated values lie, it is worthwhile considering the familiar relative frequency histogram illustration of the output results. A histogram simply involves arranging the calculated values in increasing (or decreasing) order, and dividing them into intervals of appropriate width. It makes no presumptions about the nature of the results and provides an unbiased illustration of the distribution of calculated values. From an appropriate histogram, the smoothed complementary cumulative relative frequency distribution curve can be constructed. These methods of presenting data are explained and illustrated in Appendix A and were summarized in the previous section.

Histograms and complementary cumulative plots for the doses resulting from the three groups of pathways and for the total dose are presented in Figures 4.2 to 4.6. Also shown on these figures are results obtained in the more familiar deterministic fashion i.e. evaluating the models using single nominal values for all parameters. The histograms all have a similar shape, common to the distributions of many parameters encountered in environmental studies. Most of the results (dose to receptor) are clustered in a narrow range, with a few found at much higher levels. Since the dose via the air pathways in particular varies over a wide range, it is difficult to represent adequately on a linear scale. Most of the lower values are lumped into the first interval and the true shape of the curve is thus somewhat obscured. Plotting the same results on a logarithmic scale (Figure 4.5) provides further insight about the distribution of results.

Figure 4.6 illustrates the total dose to the receptor from all six pathways modelled. The typically skewed distribution is again apparent. The complementary cumulative frequency curve shown in this figure also contains

FIGURE 4·2

DOSE VIA WATER PATHWAYS



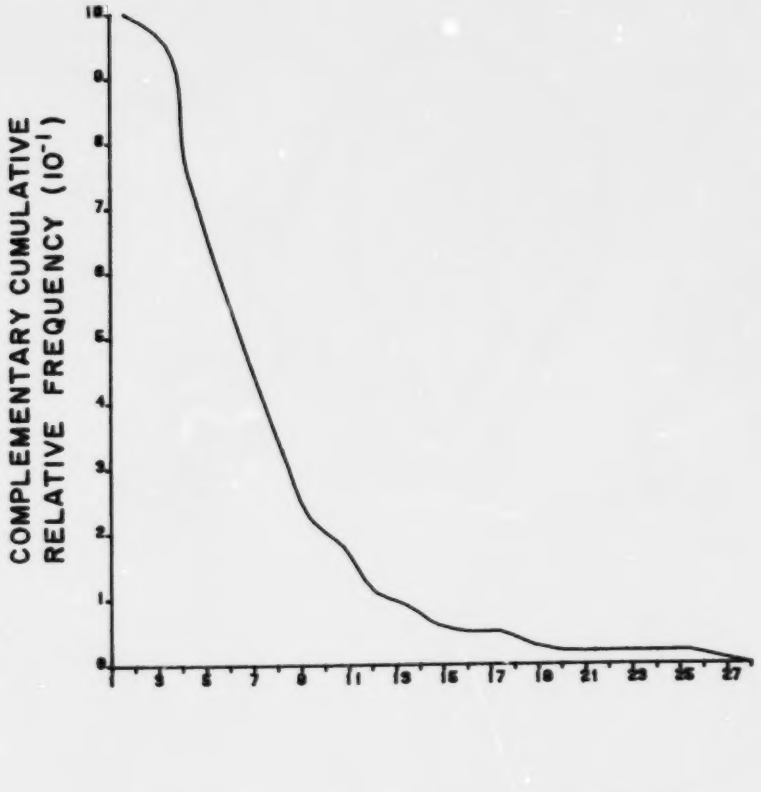
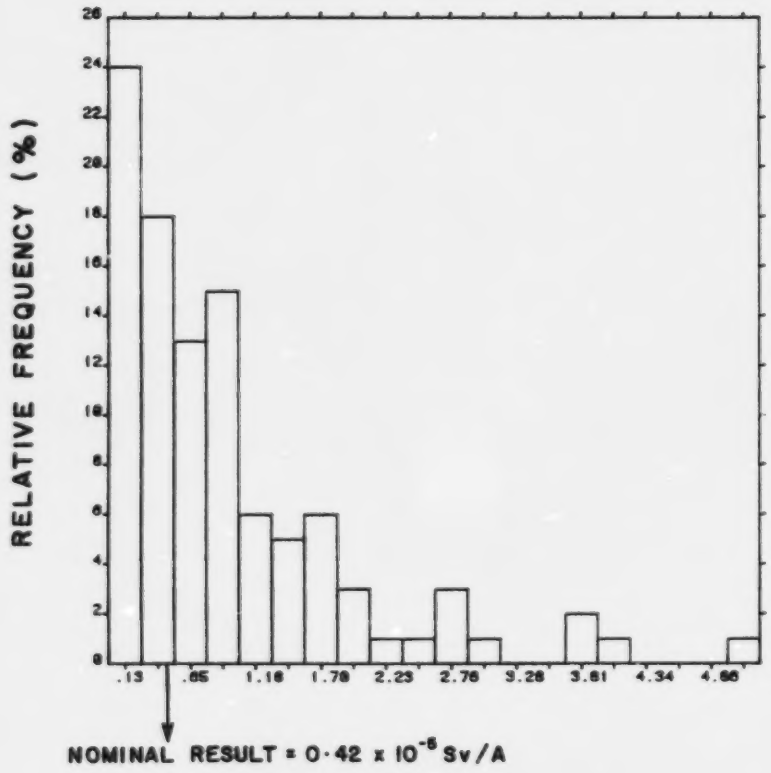


FIGURE 4·3

DOSE VIA RADON PATHWAY



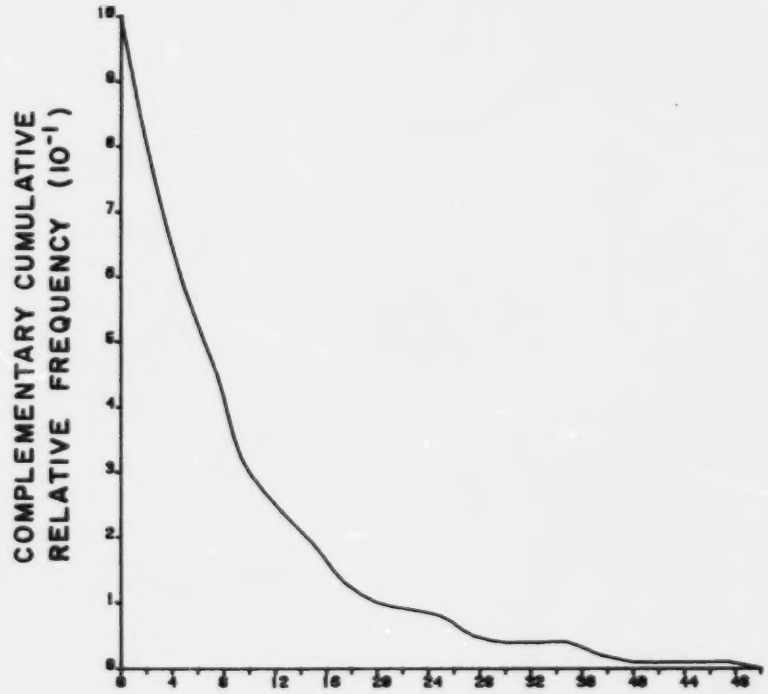
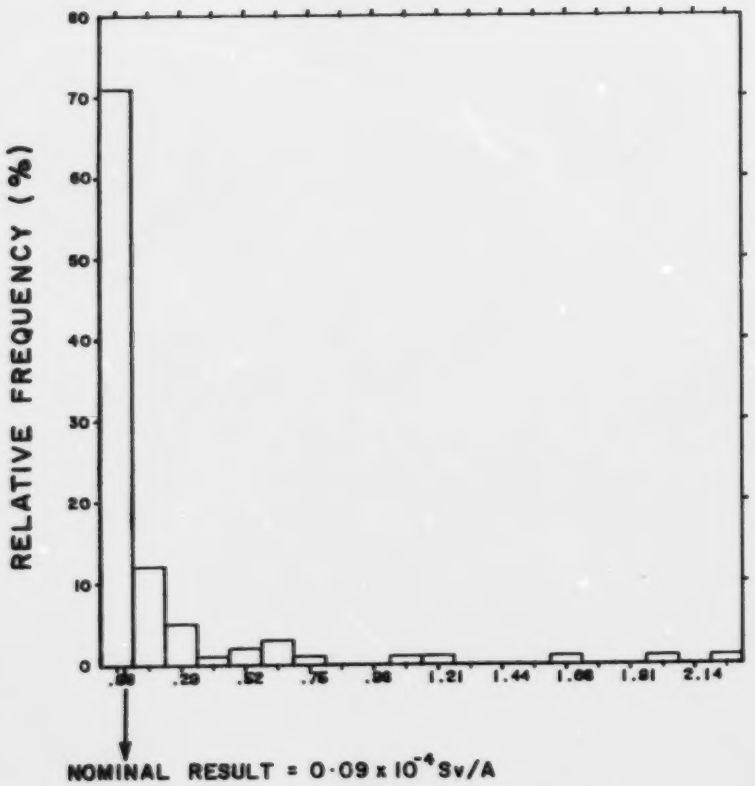


FIGURE 4·4

DOSE VIA AIR PATHWAYS



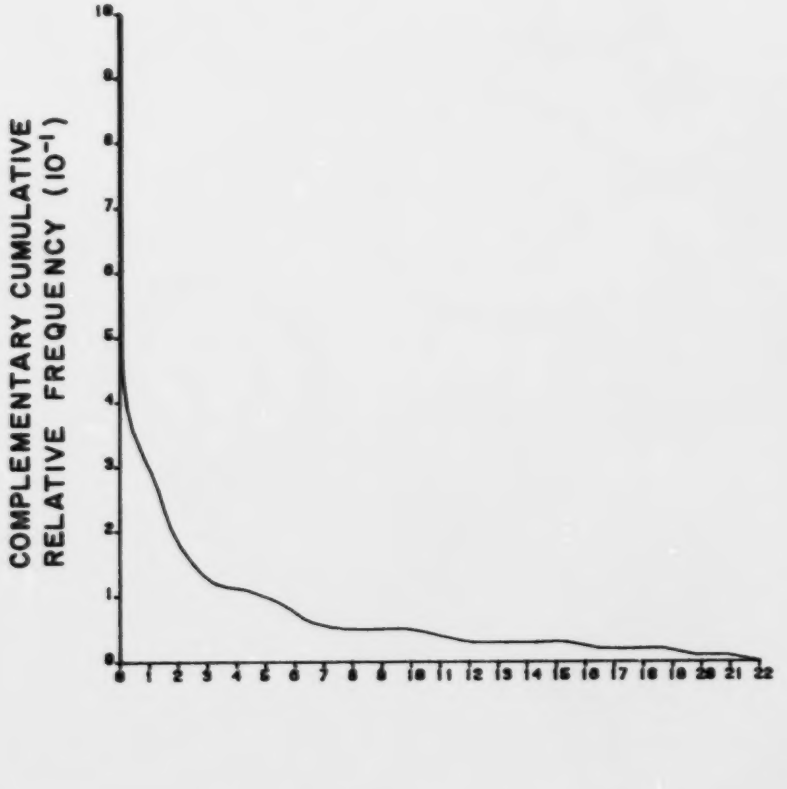
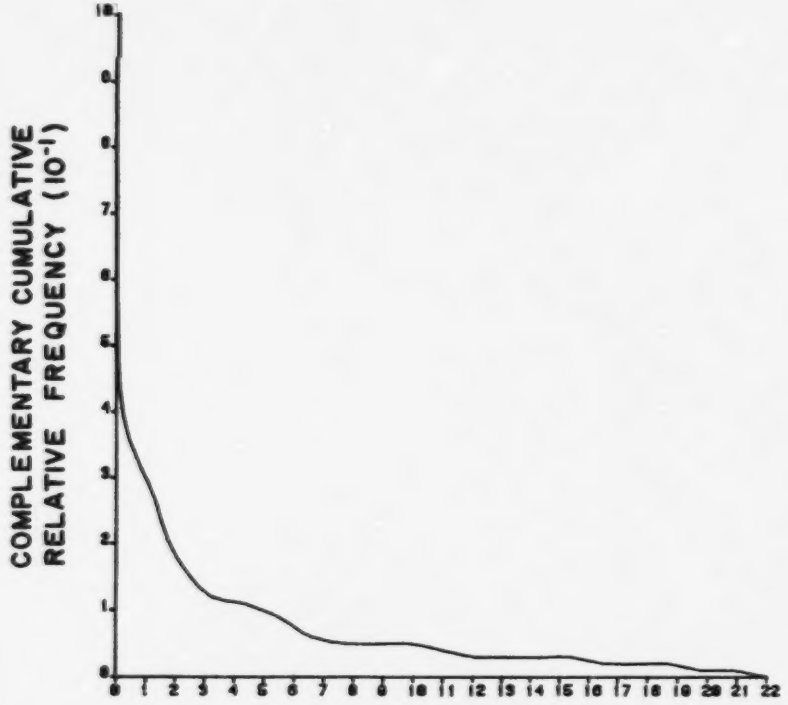


FIGURE 4·5

DOSE VIA AIR PATHWAYS
COMPARISON OF LINEAR AND LOGARITHMIC
CUMULATIVE DISTRIBUTIONS


COMPLEMENTARY CUMULATIVE
RELATIVE FREQUENCY (10^{-1})
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
DOSE VIA AIR PATHWAYS (Sv^{-6}/A)
(LINEAR SCALE)

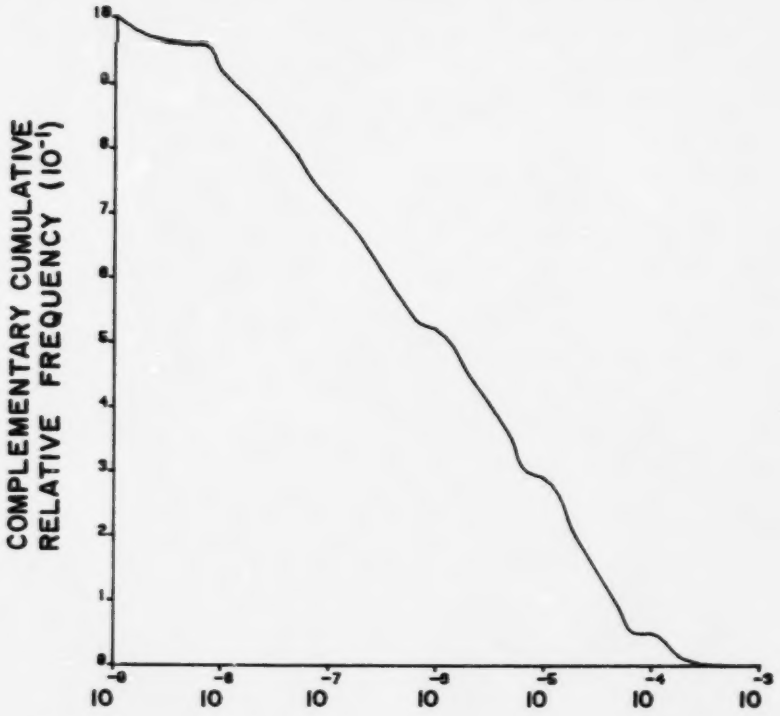
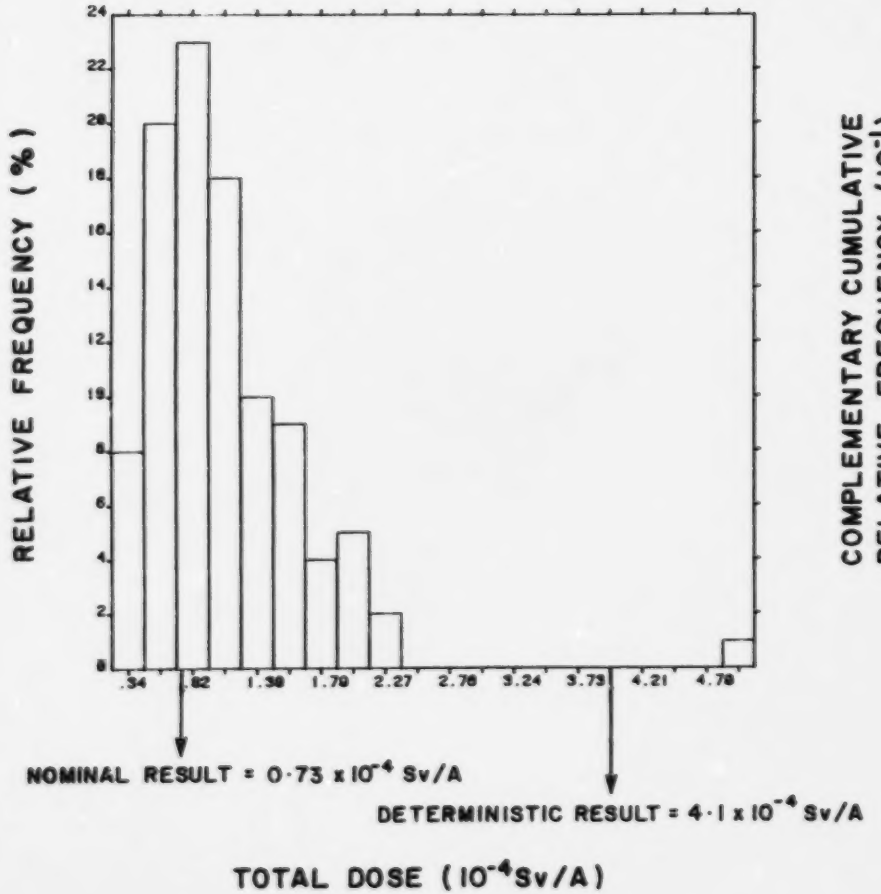
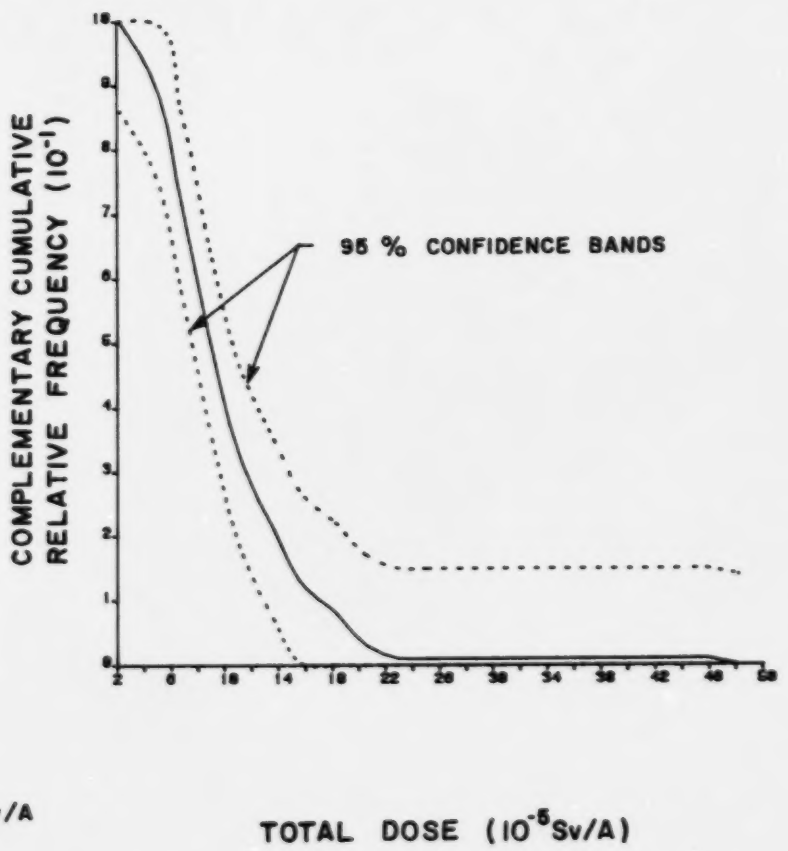


FIGURE 4·6

TOTAL DOSE TO SPECIFIED RECEPTOR





dashed curves representing the 95% confidence bands about the calculated values. Details concerning the calculation and interpretation of these 95% confidence bands can be found in Appendix A. Very briefly, they indicate the region in which the true cumulative curve is expected to lie with 95% confidence. Also shown on Figure 4.6 are two results obtained by assigning each of the input parameters a specific (i.e. single) value. The first value titled "nominal" on Figure 4.6 is the value obtained when each of the parameter values was set at its nominal 'mean' value. The second value titled "deterministic" on Figure 4.6 is the value obtained when each of the parameters was set out at nominal value plus or minus one (nominal) standard deviation (approximately) whichever direction produced the larger dose. The nominal mean value is clearly in the central part of the distribution where it would be expected. The deterministic value derived using conservative (but not tremendously so) parameter values is nearly an order of magnitude greater than the nominal value and falls in the uppermost tail of the distribution. This illustrates that compounding of even moderately conservative assumptions can lead to a high degree of overestimation in the overall result. If the deterministic result had been calculated using individual parameter values equal to the nominal values plus or minus two or more standard deviations, as is often done, then the total dose determined for the simple example presented here would have been several orders of magnitude greater than the dose estimated using nominal parameter values. With more complex models the effect of parameter value selection can be even more dramatic in calculating a single value, namely the conventional deterministic dose estimate.

Any attempt to summarize the information presented graphically in these figures results in a loss of detail. For instance, if only the mean and variance are reported, then information regarding the extreme values of the results is lost. If only extreme values are summarized (i.e. maximum and minimum) one is left with no idea of what might be a most probable value. There is further disagreement as to whether the arithmetic or geometric mean is a more valid measure of central tendency, given the shapes of the distributions. For these reasons, a variety of descriptive statistics for the four cases are provided in Table 4.1.

Table 4.1

Probabilistic Assessment Results: Summary Statistics

Source	Dose to Receptor (mSv/a) min.	max.	arith. mean	arith. SD	geom. mean	geom. SD
Radon Pathway	4.64×10^{-4}	4.97×10^{-2}	9.11×10^{-3}	9.08×10^{-3}	5.73×10^{-3}	2.78
Water Pathways	7.01×10^{-3}	4.71×10^{-1}	7.94×10^{-2}	5.50×10^{-2}	6.80×10^{-2}	1.73
Air Pathways	5.07×10^{-6}	1.49×10^{-1}	1.47×10^{-2}	2.86×10^{-2}	1.17×10^{-3}	1.81×10^{1}
All Pathways	3.29×10^{-2}	4.75×10^{-1}	1.03×10^{-1}	5.80×10^{-2}	9.16×10^{-2}	1.60

Readers familiar with the assessment of radiation in the environment from uranium mill tailings, will realize that these results are all within 'reasonable' ranges (i.e. in the order of 0.1 mSv per year or 10 mrem per year). In conclusion, it is fair to state that the probabilistic assessment tool appears to produce reasonable results based on the models used.

A further objective of this study was to illustrate that the probabilistic assessment tool could consider the effects of different tailings management options. Again, a simple example, the radon pathway model, was used to examine this question. The radon dose to the receptor was first calculated, assuming a completely uncovered tailings area. This scenario allows radon gas to diffuse unimpeded from the tailings. In the second option, the tailings were specified to be covered with a varying thickness of uncontaminated fill material. The cover thickness was specified as being triangularly distributed with a range from 0.5 m to 2 m and a mode at 1 m. This cover material would theoretically inhibit the release of radon from the tailings and should therefore result in a lower dose to the receptor.

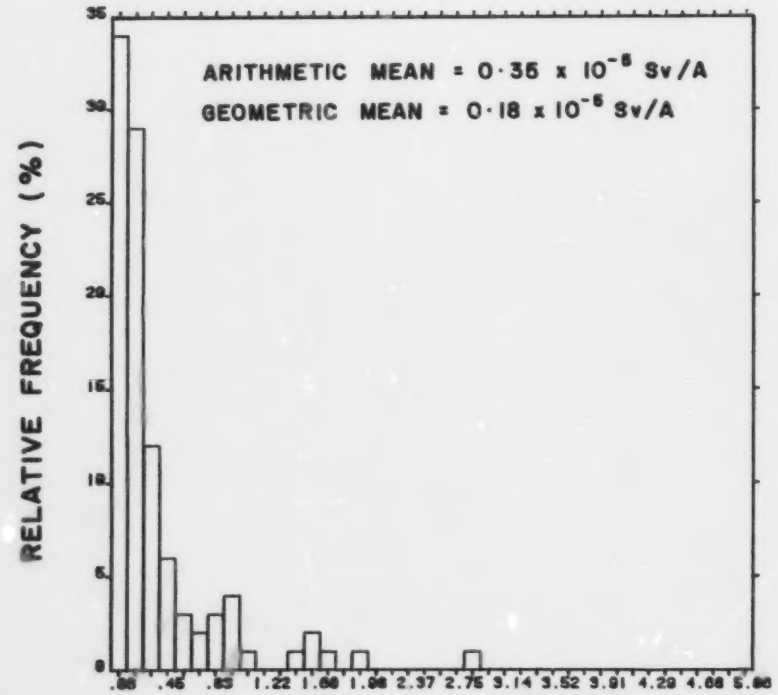
The two histograms shown in Figure 4.7 provide a graphic illustration of the effect of the two tailings management options on resultant dose. The results do show that the uncovered tailings scenario produces higher doses (arithmetic mean dose from radon of 0.35×10^{-5} $Sv.a^{-1}$) to the receptor compared to the covered tailings scenario (arithmetic mean dose from radon of 0.91×10^{-5} $Sv.a^{-1}$). This simple example provides a preliminary indication that the probabilistic assessment tool can indeed handle different tailings management scenarios.

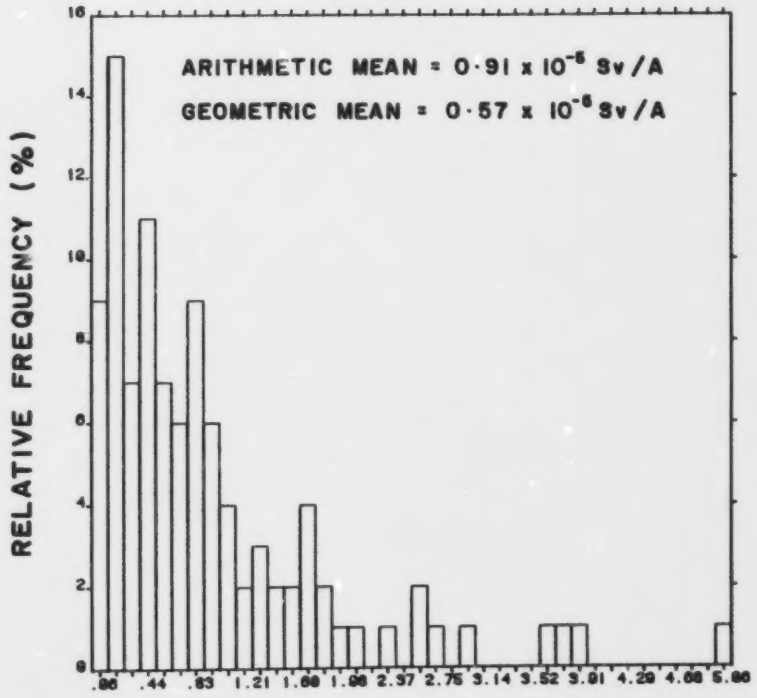
4.3 Sensitivity and Uncertainty Analysis

Although the results just presented are useful by themselves, further information is often desired. A common question is: "Which parameters/pathways have the greatest effects on model predictions?" The answer to this question may be used to direct future research work towards

FIGURE 4·7

COMPARISON OF DOSE VIA RADON FOR COVERED AND UNCOVERED TAILINGS





those parameters/pathways which are determined to be the most important. In addition, parameters and model components which are found to be relatively unimportant may be lumped or modified to simplify the analysis. A number of methods, of varying complexity, commonly referred to as sensitivity and uncertainty analysis techniques, exist to address this issue.

Before proceeding further, it is important to understand the distinction between sensitivity and uncertainty analysis. Sensitivity analysis refers to the sensitivity of the magnitude of the result to changes in the magnitudes of the input parameters. In other words, how much the mean (and distribution) move when a parameter is changed. Uncertainty analysis refers to how much the uncertainty (or spread) associated with each parameter contributes to the total uncertainty (or spread) of the result. These ideas are distinct and different. A parameter to which the value of the result is very sensitive may contribute significantly to the position of the distribution but little to its shape. Another parameter may not move the distribution very much (insensitive) but may broaden it considerably, thus adding to the 'uncertainty'. A clear appreciation of these ideas is important to the following discussion.

To further complicate matters, a number of statistical techniques exist which serve for both sensitivity and uncertainty analysis and answer the two questions simultaneously. These will be addressed briefly in Section 4.3.3.

4.3.1 Sensitivity Analysis

Classical sensitivity measures were originally defined as the rate of change of the result (i.e. dose in this study) corresponding to a change in a particular input parameter. The approach involves changing each parameter of a model by a small amount while leaving all other parameters at pre-selected nominal values. The results from this analysis essentially represent a series of partial derivatives and characterize the model response at distinct points on the response surface.

It should be noted that this method measures linear responses of the model only to small changes in input parameters. Since the parameters used in this study often vary over large ranges, the method was found to be of limited use.

A more useful measure of sensitivity is the 'sensitivity index' described by Hoffman and Gardner (1983). This method is very similar to the previous one in that individual parameters are varied while all other parameters remain at some fixed nominal value. The relative change in the model response to the change in the input parameter forms the basis for this index. Its advantage over classical methods is that the parameters are varied over their entire ranges, thus providing a more general measure of sensitivity. The sensitivity index (S.I.) is a number between 0 and 1.0 and is calculated as:

$$S.I. = 1 - \left(\frac{\text{low dose result}}{\text{high dose result}}\right)$$

Indices near 1.0 indicate parameters to which the model is very sensitive while very small numbers indicate parameters of little influence. Note that a high parameter value may sometimes result in a lower dose, depending on the functional relationship between that parameter and the calculated dose. The sensitivity indices for the various parameters in the three groups of pathways modelled are presented in Tables 4.2 to 4.4. Also indicated on these tables is an 'uncertainty index' which will be described shortly.

As a note of caution, the sensitivity index should not be used to rigorously rank parameter importance. It is most useful in indicating those parameters to which the model is very insensitive. These parameters might then be specified as constant values and/or lumped into an aggregate parameter, with little loss in model accuracy. A sensitivity index of less than 0.01 indicates numerical insensitivity of the model to the parameter. Such parameters should be given very low priority for further research or experimental characterization.

It should be noted again that the results of this study are based on

Table 4.2

Water Pathways: Radium-226 Dose Sub-Model

Sensitivity and Uncertainty Analysis Results

Parameter	Parameter Definition	Sensitivity Index	Uncertainty Index
KPRIME	Ratio of O_2 Half Saturation Constant to O_2 Conc. at Tailings Surface	0.19	1.69
NPRCP	Net Precipitation	0.78	0.317
SULPRD	Sulphate Production Rate	0.68	1.082
RHOTAL	Tailings Bulk Density	0.14	0.025
O2DIFF	Oxygen Diffusivity Through Tailings	0.49	0.318
RAGPTA	Radium-226 Content of Gypsum	0.71	0.366
SIGSS	Suspended Solids Sedimentation Velocity	<0.01	0.025
KSUBD	Radium-226 Sorption Coefficient onto Suspended Solids	<0.01	5.054
SIGBIO	Biomass Sedimentation Velocity	<0.01	0.053
KSUBBI	Radium-226 Bioaccumulation Factor by Aquatic Biota	0.17	3.948
SSR	Suspended Solids Level in River	<0.01	1.764
PROBIO	Biomass Productivity in Lake	0.03	.069
KSUBL	Radium-226 Transfer Coefficient from Sediment to Water	<0.01	2.29
ZETA	Ratio of Radium-226 in Interstitial Water to Lake Water	0.10	0.318
BFRA	Radium-226 Transfer Coefficient: Water to Fish Flesh	0.89	6.213

Table 4.3

Air Pathways: Radium-226 Dose Sub-Model

Sensitivity and Uncertainty Analysis Results

Parameter	Parameter Definition	Sensitivity Index	Uncertainty Index
RESFAC	Resuspension Factor	1.00	0.862
RHOTAL	Tailings Bulk Density	0.78	0.005
RESDEP	Resuspension Depth	0.90	0.060
HSUBD	Initial Mixing Depth	0.66	0.010
WIND	Wind Speed	<0.01	0.001
RACNTA	Radium-226 Concentration in Tailings	0.67	0.004
FRACT	Fraction of Time Wind Blows Towards Receptor	0.75	0.004
SIGZED	Vertical Dispersion Coefficient	0.96	0.022
OCCFAC	Occupancy Factor	-	0.010

Table 4.4

Radon Pathway: Dose Sub-Model

Sensitivity and Uncertainty Analysis Results

Parameter	Parameter Definition	Sensitivity Index	Uncertainty Index
MPCH2O	Moisture Content of Tailings	0.92	0.061
RACNTA	Radium Concentration in Tailings	0.67	0.067
EMANAT	Radon Emanation Coefficient	0.99	0.268
RHOTAL	Tailings Density	0.46	0.008
THICK	Clean Cover Thickness	0.78	0.061
LSUBC	Radon Diffusion Length to Cover	0.74	0.085
WTSFLX	Winter to Summer Flux Ratio	0.32	0.518
FRACT	Fraction of Time Wind Blows to Receptor	0.75	0.032
WIND	Wind Speed	0.68	0.018
SIGZED	Vertical Dispersion Coefficient	0.96	0.248
HSUBD	Initial Mixing Depth	<0.01	0.067
OCCFAC	Occupancy Factor	-	0.034

preliminary models and parameter distributions and thus should be viewed accordingly. The sensitivity and uncertainty analysis results merely illustrate the efficacy of a variety of techniques and should not be interpreted otherwise.

Although the sensitivity index is conceptually quite simple and the results correspond well with those suggested by experience, it may be misleading in some instances. The sensitivity results will vary depending on the nominal values chosen for the other parameters while examining the effects of one particular parameter. In mathematical terms, the sensitivity of a model to a given parameter will be different depending on the location on the response surface. Some of the difficulties inherent in this type of sensitivity analysis can be overcome by applying the combined sensitivity/uncertainty analysis methods outlined in Section 4.3.3.

4.3.2 Uncertainty Analysis

Uncertainty analysis, as defined earlier, attempts to quantify the contribution of each parameter distribution to the total uncertainty in the model predictions. Although a relatively simple approach to this analysis was attempted, it is not completely valid for the models used in this study.

The variance of the result of a perfectly additive or multiplicative model can be estimated by combining the variances of the input parameters. For a purely additive model;

$$Y_a = X_1 + X_2 + X_3 \ldots + X_n$$

the variance of Y_a is simply the sum of the variances of the individual parameters:

$$\mathrm{Var}(Y_a) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \mathrm{Var}(X_3) \ldots + \mathrm{Var}(X_n)$$

For a purely multiplicative model;

$$Y_m = X_1 * X_2 * X_3 \ldots * X_n$$

a log transform must first be performed, resulting in an additive model:

$$\ln Y_m = \ln X_1 + \ln X_2 + \ln X_3 + \ldots + \ln X_n$$

The variance of $\ln Y_m$ is equal to:

$$\mathrm{Var}(\ln Y_m) = \mathrm{Var}(\ln X_1) + \mathrm{Var}(\ln X_2) + \mathrm{Var}(\ln X_3) + \ldots + \mathrm{Var}(\ln X_n)$$

which is easily determined from the geometric standard deviation (GSD):

$$\mathrm{Var}\ (\ln Y_m) = [\ln\ (\mathrm{GSD}(Y_m))]^2$$

Since variance is a reasonable measure of parameter uncertainty, the ratio of the variance of a particular parameter to the overall variance of the result provides an estimate of the contribution of that parameter to model uncertainty. A ratio using this concept has been termed the 'importance index' (Hoffman and Gardner, 1983). However, preference has been given in this study to the use of the term "uncertainty index" to maintain consistency with the earlier definition of uncertainty analysis.

For an additive model (as described earlier), the uncertainty index (U_i) is given by:

$$U_i = \frac{\mathrm{var}\ (X_i)}{\mathrm{var}\ (Y_a)}$$

While for a multiplicative model, the uncertainty index is calculated by:

$$U_i = \frac{\mathrm{var}\ (\ln X_i)}{\mathrm{var}\ (\ln Y_m)} = \frac{[\ln(\mathrm{GSD}(X_i))]^2}{[\ln(\mathrm{GSD}(Y_m))]^2}$$

The uncertainty indices for each of the parameters in the 3 groups of pathways implemented, are presented in Tables 4.2 to 4.4. Note that the pathways have been assumed to be multiplicative rather than additive models (i.e. U_1 calculated using geometric standard deviations).

Unfortunately, this index is only strictly valid for purely additive or multiplicative models. For such models, the parameter uncertainty indices would all be less than 1.0 and would also add to 1.0. A parameter uncertainty index near 1.0 would indicate that the particular parameter contributed significantly to uncertainty in the result. Conversely, an index of near zero indicates little contribution to total variation in the result.

The results in Tables 4.2 to 4.4 illustrate the deviation of the models used from those of a purely multiplicative nature. The uncertainty indices presented are likely of little practical use. However, the total dose to the receptor is derived by simply adding the doses resulting from the three groups of pathways. Since the calculation is purely additive; i.e.

$$\begin{pmatrix}\text{dose to}\\ \text{receptor}\end{pmatrix} = \begin{pmatrix}\text{dose via}\\ \text{radon}\end{pmatrix} + \begin{pmatrix}\text{dose via}\\ \text{water}\end{pmatrix} + \begin{pmatrix}\text{dose via}\\ \text{air}\end{pmatrix}$$

the relative importance of the three groups of pathways can be examined using the uncertainty index. This is summarized in Table 4.5.

These results provide a preliminary indication of pathway contributions to total uncertainty. For the particular example assessed in this study,the uncertainty index indicates that the water pathways are probably more significant than pathways with an air source term (radium or radon). Again, no firm conclusion should be drawn from these very preliminary results.

Table 4.5

Contribution of Pathways to Total Uncertainty

	$\bar{X}_i$	SD_i	U_i
Radon	.91E-5	.91E-5	.02
Air	.15E-4	.29E-4	.21
Water	.79E-4	.55E-4	.77
Total	.10E-3	.58E-4	1.0

where

$\bar{X}_i$ = arithmetic mean of results

SD_i = standard deviation of results

U_i = uncertainty index = $\frac{Var_i}{\sum Var_i} = \frac{(SD_i)^2}{\sum [(SD_i)^2]}$

4.3.3 Combined Sensitivity/Uncertainty Analysis

Several techniques are available to address the issues of sensitivity and uncertainty analysis simultaneously. One is a relatively simple technique which produces a very graphic illustration of parameter importance, while the others are based on statistical correlation and regression theory.

The first technique is the SENSYV approach outlined in an AECL report by Wuschke et al (1983). In this approach a parameter of interest is held constant at a chosen value in its range while all other parameters are allowed to vary within their specified distributions. Typically, the parameter under examination is held constant at the maximum and minimum values of its distribution. The results of these 'high' and 'low' runs are compared with those of a reference case (in which all parameters are varied) by graphically superimposing the cumulative plots. Selected statistics of the dose result can also be compared. This provides a comparison between effects of different levels of variables given the range of variation that exists in the remaining variables.

Three parameters were selected from the radon pathway model to illustrate this technique. Results for these three parameters (tailings moisture, the radon emanation coefficient and the tailings bulk density) are given in Figures 4.8 to 4.10. The noticeable spread in the curves on Figure 4.8 and 4.9 suggest that the model is 'sensitive' to changes in either tailings moisture or the emanation coefficient but relatively insensitive to changes in the tailings bulk density (Figure 4.10). These indications correspond well to those provided by the sensitivity index. One drawback to this method is that it requires a large amount of computer time to perform the sensitivity/uncertainty analysis on all model parameters. However, the few results presented here indicate that the method provides a general, easily understandable measure of parameter importance.

Summary statistics on the sensitivity/uncertainty analysis results shown on Figures 4.8 to 4.10 are presented on Table 4.6. The statistics given on the

FIGURE 4·8

EFFECT OF MPCH2O
(MOISTURE CONTENT OF TAILINGS)

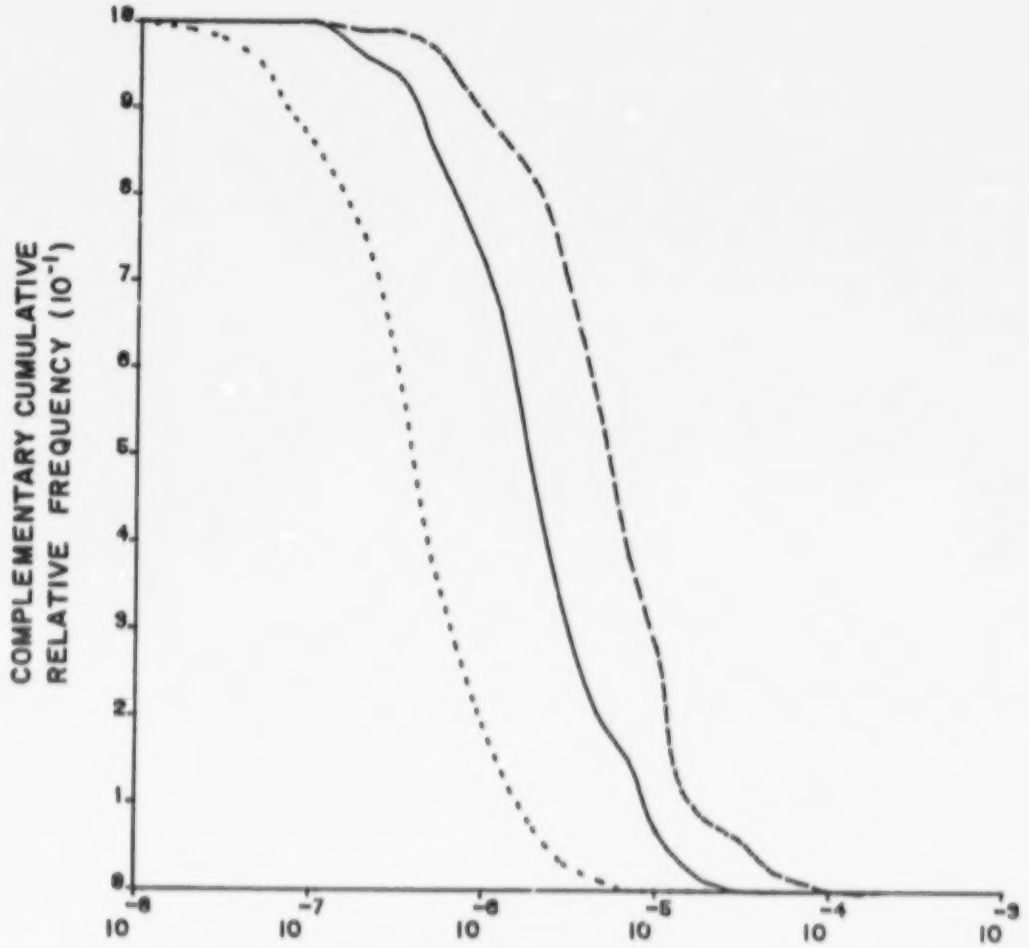


FIGURE 4·9

EFFECT OF EMANAT (EMANATION COEFFICIENT)

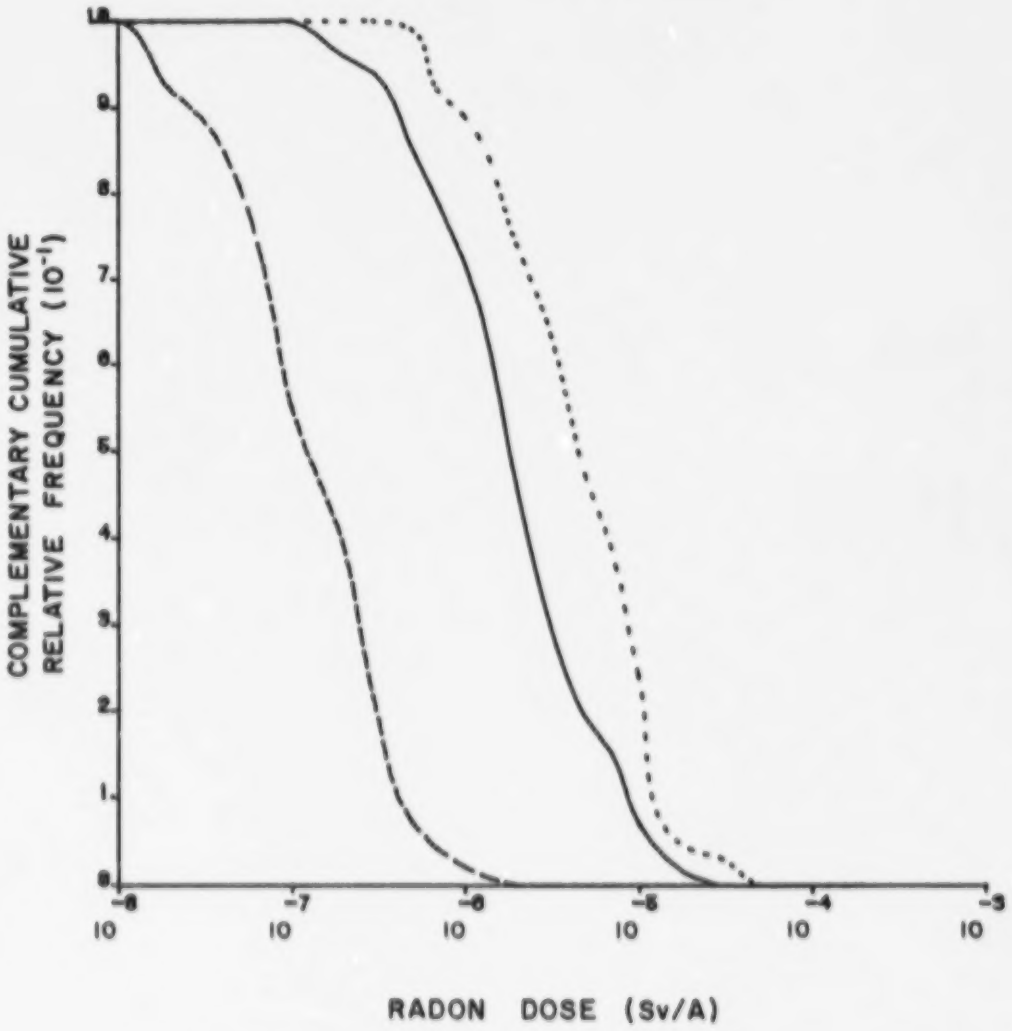


FIGURE 4·10

EFFECT OF RHOTAL
(TAILINGS DENSITY)

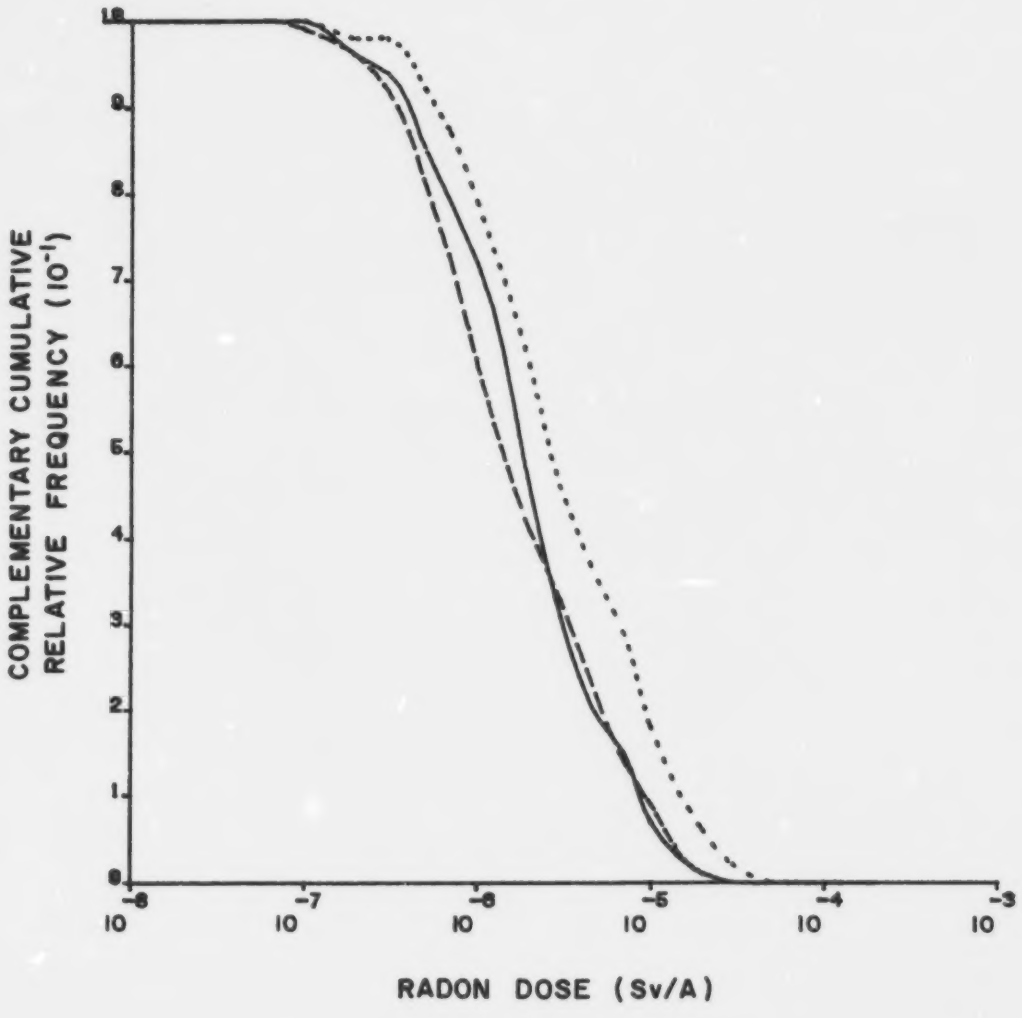


Table 4.6

Combined Sensitivity/Uncertainty Analysis (TR-206 Method)

Radon Dose to the Receptor ($Sv.a^{-1}$)

Case	Mean ($\overline{X}$)	Std. dev. (SD)	z	F
reference (all variable)	.35E-5	.46E-5	-	-
MPCH20 Low	.93E-5	.13E-4	4.2	7.8
MPCH20 High	.68E-6	.95E-6	5.9	2.3E1
EMANAT Low	.20E-6	.22E-6	7.1	4.4E2
EMANAT High	.73E-5	.79E-5	4.2	2.9
RHOTAL Low	.32E-5	.40E-5	4.7E-1*	1.3*
RHOTAL High	.60E-5	.74E-5	2.8	2.6

Notes:

MPCH20 = % moisture in tailings
EMANAT = radon emanation coefficient
RHOTAL = tailings density (g/m^3)

$$z = \left(\frac{\overline{X}_1 - \overline{X}_2}{SE} \right) \text{ where } SE = \sqrt{\frac{2SD^2}{n}} \text{ and } SD^2 = \left(\frac{SD_1^2 + SD_2^2}{2} \right)$$

$$F = \frac{SD_1^2 \text{ (larger variance)}}{SD_2^2 \text{ (smaller variance)}}$$

* no significant difference between reference and test case where at the 95% level $z = 1.96$ and $F = 1.40$
99% level $z = 2.60$ and $F = 1.50$

table include the calculated arithmetic mean and standard deviation of the receptor dose for seven separate runs consisting of 100 trials each. Also included on the table are the standard normal variate (z statistic) and the F statistic. These provide a measure of the statistical difference between the mean and variance of the reference case and the mean and variance of each of the test cases. The statistics shown indicate that the tailings density (RHOTAL) has the least effect on the dose estimates of the three parameters tested and that there is no significant difference (at the 99% confidence level) between the means and variances of the reference case and the case where the tailings density is at its minimum value. The summary statistics largely support the intuitive conclusions which can be drawn from Figures 4.8 to 4.10.

The recent trend in sensitivity/uncertainty analysis has been toward using statistical regression techniques to determine parameter importance. These are often referred to as 'statistical sensitivity analyses'. A discussion of multiple regression analysis is beyond the scope of this report, and the interested reader is referred to one of the many available texts on the topic. In brief, a state of the art statistical computer package, such as rank regression using PRESS (Iman et al 1980), will provide a wide spectrum of analysis options. The results of any type of multiple regression may be interpreted in several ways to establish valid rankings of parameter importance. The squared value of the simple correlation coefficient (r^2) provides a direct estimate of the percentage of variability in the model prediction attributable to variability in individual parameters.

A multiple regression analysis can be performed either on the raw data produced by a probabilistic assessment or on various transformations (e.g. rank, logarithmic) of the data. Regression is usually done initially on the raw data and a preliminary parameter ranking established. If the predictive ability of this regression is low (indicated by low overall r^2) the preliminary rankings may be misleading and regressions on transformed data are often performed. Regression on the ranks of the data has been shown to provide significant improvement in predictive ability for data showing

monotonic but not necessarily linear behaviour. The parameter ranking established by the regression method showing the greatest predictive ability can be used to guide future research directions.

In summary, the sensitivity/uncertainty analyses performed in this project indicate that a variety of approaches are available. In general terms, it is not possible to rate one method as 'best' or superior to any other. Each approach has its own strengths and weaknesses and each application its own requirements. The appropriate technique must be selected for each application bearing in mind the issues to be addressed and the end use of the results. The material presented here does show that results produced by the probabilistic assessment tool can be used to perform a wide variety of sensitivity/uncertainty analyses. Further discussion of this topic is provided in Appendices E.2 and E.3.

4.4 Consideration of Transient Phenomena

Modelling of the long-term effects of uranium tailings implicitly requires that time dependency must be taken into account to reflect changes which may occur in the physical nature of a tailings site. Time dependent phenomena may include processes such as the erosion of cover material, the gradual change in radium porewater concentration through pyrite oxidation, depletion of gypsum and other phenomena which tend to occur steadily over time. The effect of episodic phenomena such as forest fires or earthquakes which are essentially random events, with a known average frequency of occurrence, could also be considered.

A variety of techniques are available to simulate time dependent phenomena. In the simplest situation the mathematical model which describes a real system may be solved analytically.

For some complicated systems, response function techniques and various numeric methods can be used to solve the system equations. Taking into account system perturbations arising from episodic phenomena is still more complex.

Situations can also be envisaged where not only is the mathematical model time dependent in some fashion, but the parameters incorporated in the model are also time dependent. These possibilities are raised only to suggest that the modelling of time dependent phenomena can be quite complex and time consuming particularly as the complexity of the simulation model increases.

For the purposes of this Phase I study the ability to incorporate simple time dependency was investigated using the groundwater pathway. Radium transport in the groundwater system was modelled taking into account the time dependent influence of adsorption onto the geologic material, and radioactive decay. The analysis was simplified by assuming a steady-state continuous source of radium input to the groundwater system and steady-state groundwater flow. The detailed derivation and documentation of this model can be found in Appendix C.

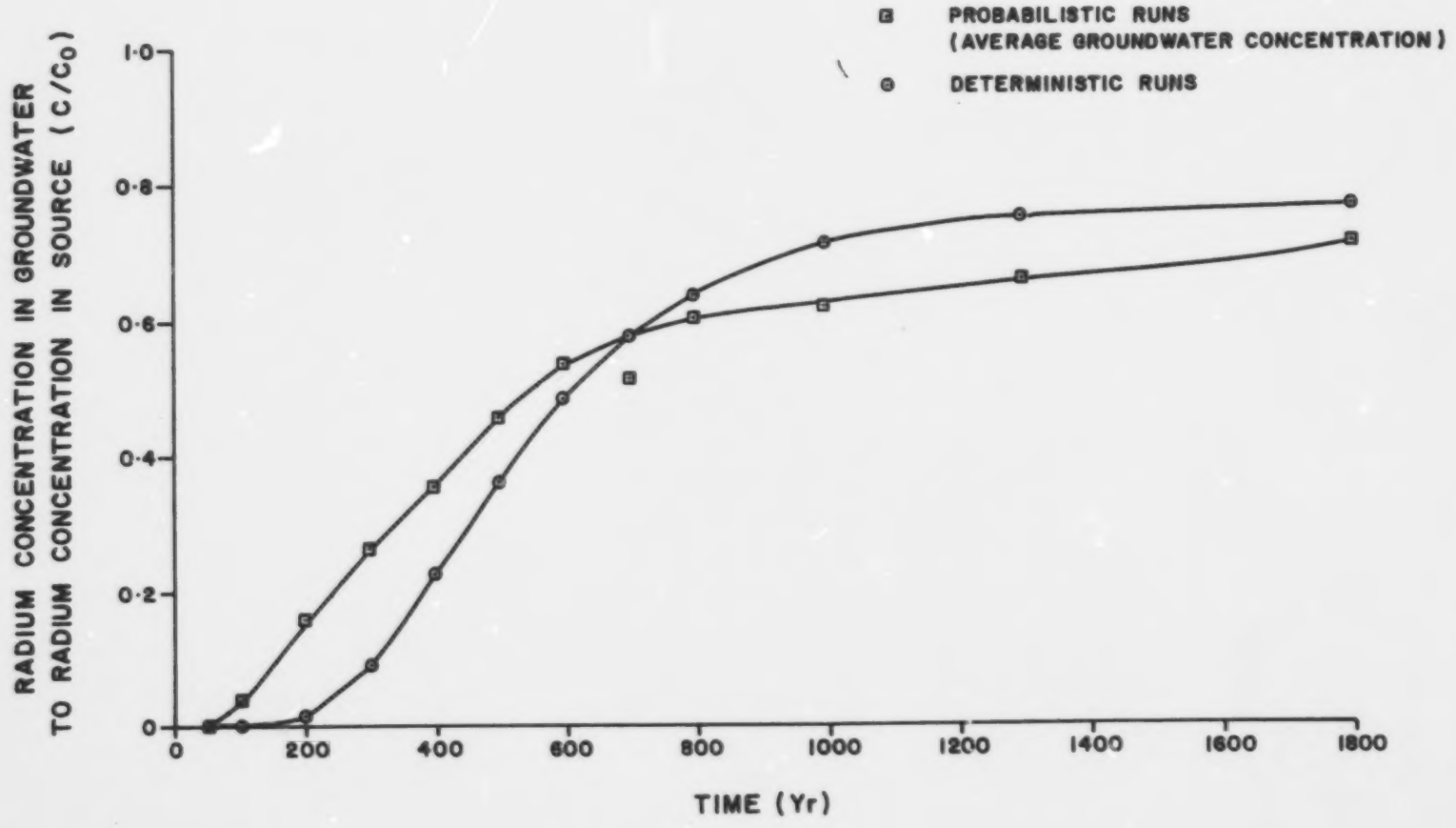
The results of a series of successive runs showed that between 150 and 200 years were required before the radium concentration began to increase in groundwater at its point of discharge to the river some 200 m away from the reference tailings site.

The variation in the groundwater radium concentration with time was examined by a series of 13 model runs for times ranging from 50 to 2500 years. The form and values of the distributions progressively changed with time approaching a quasi-steady-state after about 1500 years. This steady-state is attained within a similar time using a deterministic modelling approach as demonstrated by the plot of mean values on Figure 4.11. This figure illustrates how the concentration of radium varies with time in the emerging groundwater relative to the constant concentration of radium in the tailings porewater (C/C_o).

These calculations demonstrate the feasibility of incorporating simple time dependency and clearly illustrate the fact that different components of the system follow different time courses. While the groundwater system requires hundreds of years to approach quasi-steady conditions releases to the surface

FIGURE 4·11

VARIATION IN THE GROUNDWATER RADIUM CONCENTRATION WITH TIME AT POINT OF DISCHARGE



water reach quasi-steady state conditions within tens of years.

4.5 Interpretation of Probabilistic Results

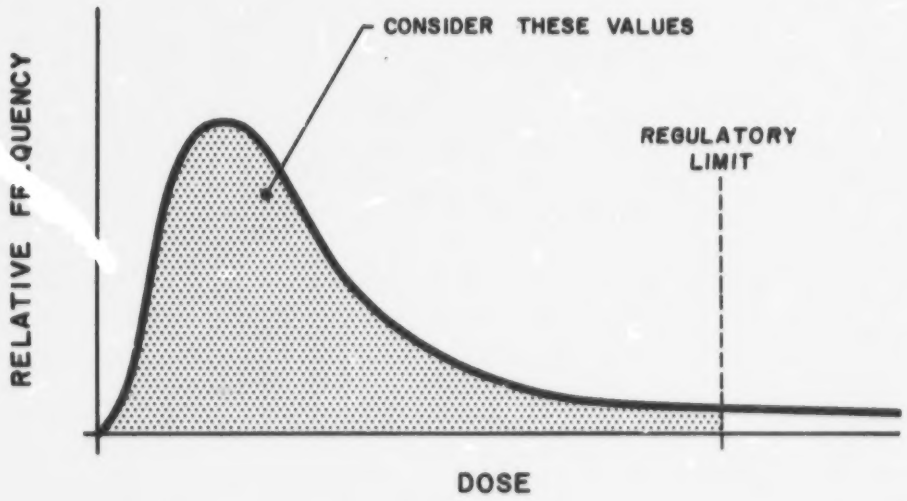
Previous sections of this report have described the probabilistic assessment concept, the particular mathematical models employed and the results obtained in this study. One area of particular importance which has not been addressed is the possible interpretations which can be applied to the results and their final use. The results of primary interest in this Phase I study have been the predicted distributions of the dose values to a specified individual. These were generally displayed as frequency histograms with high left hand peaks and long right hand tails. Given the nature of the models used, the results considered could just as easily have been distributions of the concentrations of various chemical species at the receptor site. Both of these quantities (concentration and dose) can be represented in similar ways, are likely subject to similar types of regulations and thus are amenable to similar interpretations. Several distinctly different approaches to the interpretation of the results have been proposed. However, no consensus has yet been reached as to which, if any, is the most appropriate.

Most current radiation or environmental protection criteria or standards are single values, usually a maximum permissible concentration, dose or activity. Appearing as a perpendicular line on a dose distribution graph, its interpretation is clear; doses should not exceed the limits (that is, dose estimates should be to the left of the line). Yet the tail of the dose distribution almost certainly can be extended (to the right) to exceed any criteria or standard, albeit with a very low probability.

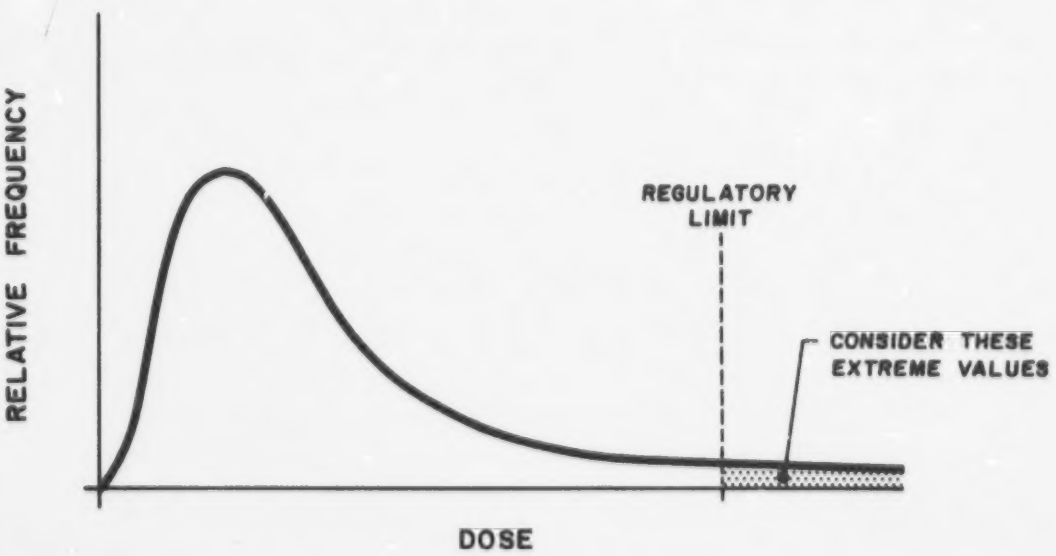
Some would argue that the tail represents situations so unlikely as to be incredible and therefore should be ignored. Others counter by asserting that these potentially high exposures should be the prime focus of regulators and that, to the individual involved, there is little consolation in knowing that the exposure was ignored because it was considered to be statistically insignificant. These two interpretations are outlined in Figure 4.12.

FIGURE 4-12

POSSIBLE INTERPRETATIONS OF DOSE DISTRIBUTIONS



A) CONSIDER BULK OF VALUES (IGNORE EXTREMES)



B) CONSIDER EXTREME VALUES (IGNORE BULK)

Advocates of the first approach (those choosing to ignore extreme values) would claim that the many sources of model uncertainty (parameter and model bias, parameter variability, etc.) preclude making decisions based on extreme or unlikely values. Hoffman and Gardner (1984) in their topical paper (Appendix E.3) argue that "meaningful estimates of maximum or extreme values in the tails of predicted distributions are currently beyond the capabilities of extant models and data sets." They find it most productive to focus attention on a "probable range" of predicted values, since this is a range that is reasonably likely to encompass a "true" value. They would go as far as to claim that, given the nature of the models and input parameter distributions, the extreme values have no physical significance and do not represent 'real cases'.

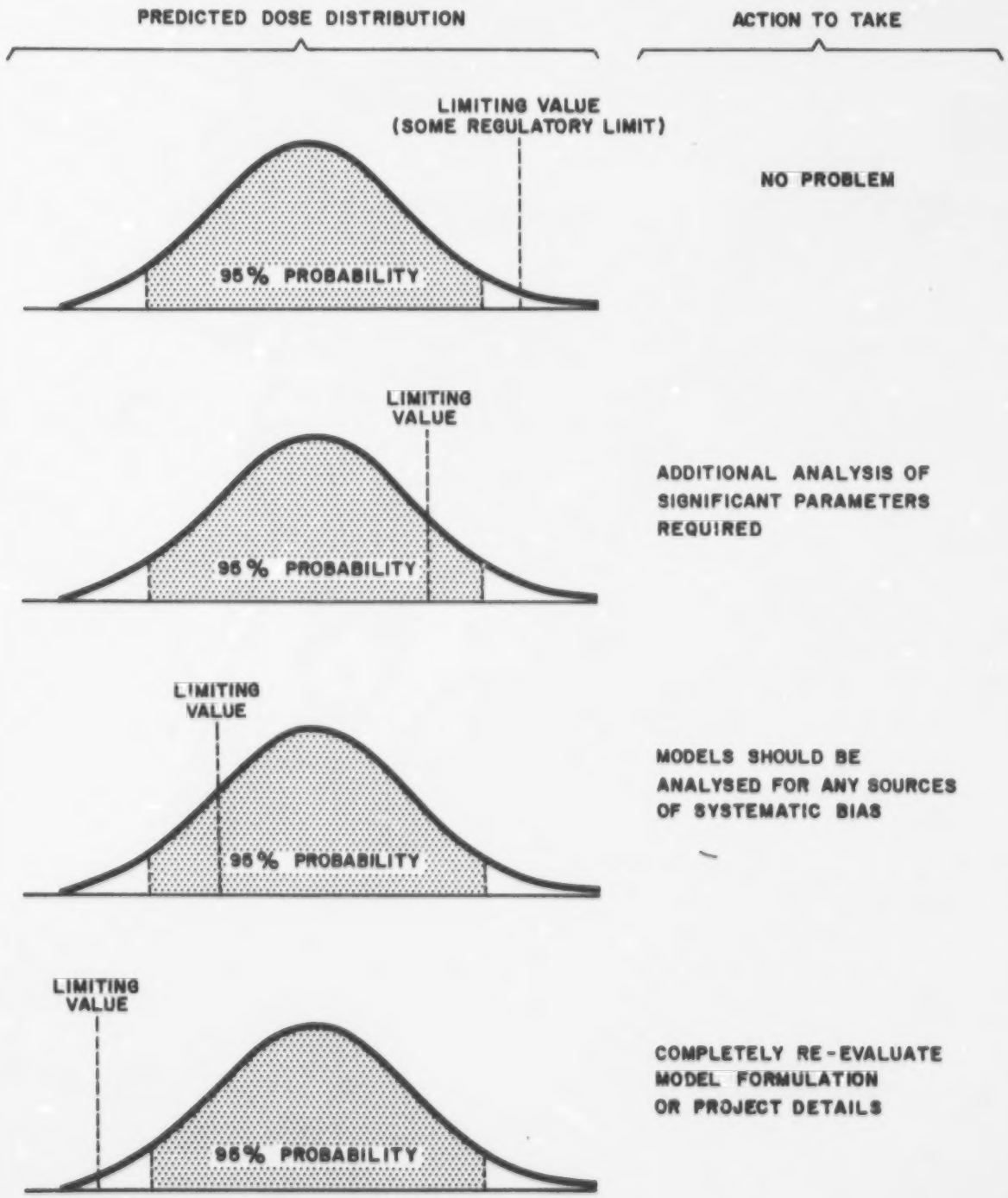
As described previously a distribution can be characterized through various statistics such as arithmetic and geometric means and variances and/or various quantiles (i.e. deciles, percentiles, etc). These can be conveniently used to arrive at an estimate of some "most probable" range of values. Typically estimates of the ranges of values encompassing 95 or 99% of the area of a distribution are of interest.

Figure 4.13 illustrates the application of these ideas. An area of the dose distribution curve containing 95% of the predicted values is characterized on the figure as being contained within the bounds of the rectangular box. The location of this area relative to a regulatory limit is shown on four successive plots. Interpretation/analysis of this type is necessarily somewhat qualitative, but may be most appropriate, given the arguments just presented. It should also be noted that this approach requires only a relatively small number of probabilistic trials. Typically, fifty to one hundred trials are sufficient to adequately estimate the necessary statistics. However, many more trials may be needed to accurately estimate the percentiles of a distribution (i.e. the tails).

Proponents of alternative approaches would argue that the extreme values of a

FIGURE 4-13

INTERPRETATION OF DOSE DISTRIBUTION (EXTREME VALUES IGNORED)



dose distribution are very much of interest to the exposed individuals and to regulators. They also point out that because of the nature of the technique, the exact shape of the dose distribution curve is unknown and variable. Because of probabilistic sampling error, the shape of the curve will change slightly with each set of trials, producing a so-called "shimmer effect". Thus, there is further uncertainty in the quantification of the likelihood of extreme values. The extreme results may in fact be considerably more or less likely than indicated by any single dose distribution curve. The level of this uncertainty can be quantified by constructing confidence bounds about the empirical curve.

Given the finite probability of these extreme events occurring and the uncertainty regarding this probability, the proponents of this second interpretation argue that the entire distribution must be considered. It has also been suggested that statistical methods be used to further predict the distribution of extreme values. Theoretical curves can be fit to the extreme values in order to predict where further extreme values might lie. A second suggestion is that the combinations of random events which lead to extreme dose predictions be considered in more detail. The probability of these scenarios may then be reduced or eliminated by considering engineering or siting changes to the system being analyzed. In this sense, extreme dose predictions may contain significant information and should not be ignored.

The entire discussion regarding the appropriate interpretation of probabilistic results is not a recent problem. Fultyn outlined the difficulties quite eloquently in 1968:

> "A probabilistic analysis allows one to examine more explictly, and hopefully more objectively, the degree of risk associated with a situation. It may also make decision-making somewhat more difficult. A traditional analysis presented one with a single, hard and fast value for an effect. That value could be compared to a single, hard and fast guideline. If the value were below the guideline, one could proceed with some confidence. If it were above the guideline, one could either not

proceed, or proceed only with some unknown peril. If a probabilistic statement of the evaluation reveals that perhaps 99.999 percent of the values of the effect can be expected to lie below a guideline value, under most circumstances one can feel secure in proceeding with the endeavor. Conversely, if a majority of the values of the effect can be expected to lie in a forbidden region, one readily realizes that a great deal of risk will be involved in proceeding with the original concept. The difficult decisions arise when perhaps 10 percent, or even only 1 percent, of the expected effects may occur above a guideline value".

The problem is obviously not yet resolved and will require additional study and discussion before a consensus is reached. To further complicate matters, several other interpretations have been proposed for application to probabilistic results. These focus primarily on a combination of the consequence of a result (eg. dose to man) and the probability of that result occurring. The mathematical combination of these quantities produces an estimate of the risk of a given situation. Risk analysis seems to be gaining increasing acceptance in evaluating environmental effects (eg. Kaplan and Garrick, 1981). An extension of this general risk-based interpretation of results attempts to consider different aspects of risk separately rather than as a single risk limit. The details of this method, which results in a complementary cumulative distribution frequency curve are not appropriate for this paper but are described in a recent reference (Cox and Baybutt, 1982). The entire issue of interpretation and ultimate use (i.e. the endpoint) of probabilistic analyses is considered in more detail in Appendix E.1.

5.0 CONCLUSIONS AND RECOMMENDATIONS

5.1 Highlights of Phase I Assessment

Before summarizing the work done in this study, it may be useful to recall the original, objective of this "Phase I" project. The study was conducted to demonstrate the applicability/feasibility of using a probabilistic assessment technique to evaluate the long-term environmental effects of uranium mill tailings. The intent was to use readily available technology to address a number of specific issues and focus on the techniques rather than the specific sub-models. It is submitted that the results of the Phase I investigations and a "first-cut" probabilistic assessment on a generic reference tailings site have demonstrated that the technique is indeed useful in this evaluation.

A review of the literature indicated that the probabilistic assessment concept is far from a new idea. References cited in the literature note that the underlying concepts of Monte Carlo simulation have been in use since about the turn of the century (Meyer, 1956). The relatively recent proliferation of inexpensive, high speed computers has led to the application of Monte Carlo and alternative probabilistic type analyses in a wide variety of fields. Early applications in nuclear research and applied mathematics have been followed by new developments in social science, business and engineering. Recently, probabilistic assessment has been used successfully in a wide variety of environmental and radioecological problems.

The review of probabilistic assessment applications indicated that a variety of computer programs are currently available for this type of analysis. A number of these were examined and some were found to be not easily transportable, unduly complex and/or unavailable. On the other hand the OPANAL program (Reilly, 1969) was found to be relatively simple, and easy to implement. Since it has been used successfully for many applications over the last 15 years and was also readily available, it was selected for the initial demonstration of the probabilistic approach.

This computer program was implemented and tested extensively to ensure that the probabilistic results produced were a function of only the input parameter distributions and the specified mathematical models and not simply artefacts of the OPANAL code. One interesting result of the test was an indication of the effect of the number of trials on the results produced. It was shown that typically 100 trials are more than sufficient to provide a stable estimate of the mean and variance of either an input or calculated model parameter. However, many more trials are required to provide an accurate estimate of the true, underlying shape of a parameter distribution. The actual number of trials required to define this shape sufficiently will depend on the anticipated final interpretation of the results which is still very much a matter of debate.

The demonstration of the concept also requires a set of mathematical models to describe the source release mechanisms and the environmental effects of contaminants at a hypothetical uranium tailings site. A reasonable, simple, yet representative scenario was defined and the appropriate environmental pathways models were formulated for six key pathways of exposure to a representative receptor. The scenario and models used were kept as simple as possible consistent with the preliminary nature of the study. The reader is cautioned that many potentially significant effects have been set aside for the moment, and the numerical results must be viewed with this in mind. While this study has demonstrated the potential usefulness of the technique, any specific application would require that considerable further attention be given to the sub-models and parameter distributions. It should be noted that increasing the sub-model complexity however, will not necessarily improve the accuracy of the results.

Despite the simplifications, it is believed that the models exercised all salient features of the probabilistic assessment technique. To further ensure that the technique was exercised fully, approximately thirty input parameters were assigned a variety of probability distributions. These distributions were selected using various combinations of experimental data, fundamental theory and professional judgement and included all parametric distributions

inherent to the OPANAL program.

The results produced by the probabilistic evaluation of the models using the OPANAL program showed trends common to environmental assessments and illustrated the flexibility of the approach. They indicated that probabilistic assessment could incorporate simple time dependency and also evaluate the effects of different tailings management options. The result of primary interest in this study however was considered to be the distribution of radiation dose to a specified individual (i.e. representative receptor). The distribution of the predicted results showed a slightly skewed (lognormal) shape, common in radioecology. Perhaps most importantly, the results produced by the probabilistic assessment were found to be suitable for a variety of interpretations and applications.

Several sensitivity and uncertainty analysis techniques were examined and applied to selected assessment runs. Sensitivity analysis, which measures the magnitude of the change in predicted results to specified changes in the magnitude of the input parameters, was found to be readily applied.

Uncertainty analysis, involving the study of the effects of parameter uncertainty (spread) on the dispersion (uncertainty) in the model results was found to be somewhat more difficult. Two unified approaches to sensitivity/uncertainty analysis were also briefly investigated. It is tentatively concluded from these investigations that there is not a single "best" method of sensitivity/uncertainty analysis. Each approach was found to have its own strengths and weaknesses and each application its own requirements.

As the study progressed, several areas requiring more in-depth analysis were revealed. A number of topical papers were prepared to increase the understanding of these conceptual and technical issues. These papers represent a synopsis of the current state of knowledge in the respective fields of study and should serve to focus further discussion in these areas.

In conclusion, the results of this study indicate that probabilistic assessment is a useful tool for the evaluation of the long-term environmental effects of uranium mill tailings. Results produced using this technique can be used to provide many different measures of model sensitivity and uncertainty. The results can also be analyzed and interpreted in a variety of ways depending on the ultimate "end point" of the assessment. The technique provides considerably more information than the traditional deterministic approach about the environmental consequences of the system of interest and thus should contribute to more enlightened decision making.

5.2 Recommendations for Further Work in Phase II

The Phase I study indicated that the probabilistic assessment technique could indeed be used effectively to examine the long-term environmental effects of uranium mill tailings. An expansion of the scope of the investigations is thus recommended for the Phase II studies. In particular, the emphasis in Phase II should initially be placed on incorporating additional contaminants (both radioactive and non-radioactive), more complete pathways models and continued investigation of a number of conceptual and technical issues. A number of more specific recommendations follow.

For a probabilistic analysis to be used effectively in evaluation and assessment of the merits of various tailings management strategies the component models used must provide reasonable approximations to reality. For this reason, the component models should be expanded in the following ways:

- all primary radionuclides and chemical contaminants of importance should be included in the analysis
- additional environmental pathways potentially providing a significant contribution to the exposure of man should be modelled.
- where appropriate, methods to include time dependent parameters and episodic events should be investigated

The results produced by a probabilistic evaluation using these more complete

component models should provide reasonable estimates of the magnitude of long-term environmental effects and their associated uncertainties. Sensitivity/uncertainty analysis should be performed on these more complete results to provide quantitative estimates of important pathways and parameters. The main objective of this analysis is to identify parameters which are especially important so that a priority list can be developed to guide further research. The object of such research is to reduce the uncertainty in parameter values and thus improve the precision (narrow the range) of model predictions.

There are a number of conceptual and technical issues which must also be addressed. These include:

- further investigation of appropriate sensitivity/uncertainty analysis techniques.
- consideration of various methods of incorporating time dependency
- the merits of calculating dose to a (representative) individual (reference receptor) versus calculation of dose to the entire population
- relevant statistical concepts including sampling strategies, effects of truncating input distributions, etc.

Perhaps most importantly, although not necessarily a part of the Phase II studies, further effort should be given to determining the ultimate use and appropriate interpretation (end-point) of the results of a probabilistic assessment. This may well be the most difficult challenge as it involves far more than objective analysis or the solution of a well-defined technical problem.

6.0 REFERENCES

Cairns, J.J. and Fleming, K.N., 1977. "STADIC - A Computer Code for Combining Probability Distributions." for San Francisco Operations Office, U.S. Energy, Research and Development Administration. Contract EY-76-C-03-0167.

Chambers, D.B., Knapp, R.A., Ibbotson, B.G. and Lowe, L.M., 1982. "Overview of Tailings Management Practice." Presented at the International Conference on Radioactive Waste Management. Winnepeg, September.

Cox, D.C., and Baybutt, P., 1982. "Limit Lines for Risk." Nuclear Technology, Vol. 57, June.

Crow, E.L., Davis, F.A., and Maxfield, M.W., 1960. Statistics Manual, Dover, N.Y.

Dormuth, K.W., and Sherman, G.R., 1981. "SYVAC - A Computer Program for Assessment of Nuclear Fuel Waste Management Systems, Incorporating Parameter Variability". Atomic Energy of Canada Ltd., Whiteshell Nuclear Research Establishment, Pinawa, Manitoba. AECL Report 6814.

Fox, D.G., 1982. "Uncertainty in Air Quality Modelling." A Summary of the AMS Workshop on Quantifying and Communicating Model Uncertainty, in Bulletin of American Meteorological Society, 65, 1, January 1984.

Fultyn, R.V., 1968. "A Probabilistic Scheme for Predicting and Evaluating Hazards Resulting from Radioactive Emissions to the Atmosphere." In Environmental Surveillance in the Vicinity of Nuclear Facilities: Proceedings of a Symposium Sponsored by the Health Physics Society, W.C. Reinig (ed), Thomas, Springfield, Ill. 1970.

Gardner, R.H., 1984. Personal Communication.

Guvanasen, V.M., Wuschke, D.M. and Csullog, G.W., 1983. "Application of SYVAC to Alternative Assessments." Presented at the 16th Waste Management Information Meeting, Winnipeg, 1983.

Henrion, M., 1981. "An Overview of DEMOS." Department of Engineering and Public Policy, Carnegie-Mellon University, Pittsburgh, Penn.

Hoffman, F.O. and Gardner, R.H., 1983. "Evaluation of Uncertainties in Environmental Radiological Assessment Models." In Radiological Assessment: A Textbook on Environmental Dose Assessment, J.E. Till and H.R. Meyer (eds), U.S. Nuclear Regulatory Commission, Washington, D.C., NUREG/CR-3332, ORNL-5968.

Kaplan, S., and Garrick, B.J., 1981. "On the Quantitative Definition of Risks." Society for Risk Analysis.

Macaluso, P. 1984. "A Risky Business - An Introduction to Monte Carlo Venture Analysis." BYTE, March.

Meyer, H.A., (ed.) 1954. Symposium on Monte Carlo Methods, Wiley, N.Y., 1956.

Monserco Ltd., 1983. "Probabilistic Consequence Assessment of Hydrogen Sulphide Releases from a Heavy Water Plant - GASPROB Code Documentation." For the Atomic Energy Control Board - INFO-0102-2.

O'Neill, R.V., 1971. "Error Analysis of Ecological Models." In Radionuclides in Ecosystems. Proceedings of the Third National Symposium on Radioecology. Oak Ridge Tennessee, CONF-710501 - P2.

Reilly, P.M. and Johri, H.P., 1969. "Decision Making Through Opinion Analysis." Chemical Engineering, Apr. 7, 1969.

Schlaifer, R., 1959. "Probability and Statistics for Business Decisions - An Introduction to Managerial Economics Under Uncertainty." McGraw-Hill, New York.

Schwarz, G. and Hoffman, F.O., 1980. "Imprecisions of Dose Predictions for Radionuclides Released to the Environment: An Application of a Monte Carlo Simulation Technique." Environment International, 4, 1980.

SENES Consultants Limited, 1983. "Decommissioning of the Beaverlodge Mine/Mill Operations and Reclamation of the Site: Environmental and Radiological Pathways Analysis." Report to Eldorado Resources Limited.

Shaeffer, D.L. and Hoffman, F.O., 1979. "Uncertainties in Radiological Assessments - A Statistical Analysis of Radioiodine Transport via the Pasture-Cow-Milk Pathway." Nuclear Technology, 34, August, 1979.

Stocum, W.E., 1968. "Variability in the Parameters Used to Predict the Dose to the Thyroid from the Ingestion of I-131 in Milk." In Environmental Surveillance in the Vicinity of Nuclear Facilities: Proceedings of Symposium Sponsored by the Health Physics Society. W.C. Reinig (ed.) Thomas, Springfield, Ill. 1970.

Wimmenauer, C., 1971. "Monte-Carlo Simulation: An Overview." For the Quality Evaluation Dept; Naval Ammunitions Depot, Crane, Indiana.

Snedecor, G.W. and Cochrane, W.G., 1967. "Statistical Methods." Iowa State University Press, Ames, Iowa.

Utah State University, 1982. "Probabilistic Modelling of Tailings Embankment Designs." For the Bureau of Mines (161-82) Washington, D.C., PB83-122598.

Wuschke, D.M., Rice, A.M. and Gillespie, P.A., 1983. "Environmental Assessment of Sub-seabed Disposal of Nuclear Wastes: A Demonstration of Probabilistic Systems Analysis." Atomic Energy of Canada Ltd., Environmental and Safety Assessment Branch, Pinawa, Manitoba, AECL Report TR-206.

APPENDIX A

AN OVERVIEW OF SOME USEFUL STATISTICAL CONCEPTS

APPENDIX A - AN OVERVIEW OF SOME USEFUL STATISTICAL CONCEPTS

This appendix provides an overview of some statistical concepts useful in carrying out and interpreting "probabilistic" analyses of the type described in the report. Further discussion of such topics can be found in standard texts such as those indicated in the references to this appendix.

A.1 Random Variables

A random variable or variate is a quantity about which there is some uncertainty. In the ordinary interpretation, this is a chance quantity, usually the outcome of a physical process which cannot be exactly predicted. However, a partial description of the outcome can be made, as for example with flipping a coin or rolling a die; we know from experience in both of these cases that each possible result is equally likely. This partial description may be quantified by stating that the probability, or long-term relative frequency, of each outcome is the same; one-half for each of "heads" and "tails" with the coin; one-sixth for each of the number of dots (1, 2, ..., 6) on the die.

A.2 Probability Distributions

The assignment of a probability to <u>every one</u> of the possible values of a random variable is called a probability distribution or simply a distribution and may be expressed as a probability function which may in turn be graphed. The probability function for the number of dots which appear when two dice are rolled is:

p(2) = 1/36	p(7) = 6/36
p(3) = 2/36	p(8) = 5/36
p(4) = 3/36	p(9) = 4/36
p(5) = 4/36	p(10) = 3/36
p(6) = 5/46	p(11) = 2/36
	p(12) = 1/36

This may be summarized as:

$$p(x) = (x-1)/36 \quad \text{for } x = 2, 3, \ldots, 7$$
$$p(x) = (13-x)/36 \quad \text{for } x = 8, 9, \ldots, 12$$
$$p(x) = 0 \quad \text{otherwise}$$

The graph of this probability function is shown in Fig. A.2.1(a).

Another way of expressing a probability distribution is by means of the cumulative probability function which gives the probability that the random variable will be equal to or less than the stated value. For example, the cumulative probability function for the number of dots on rolling two dice is:

$$p(x) = 0 \quad \text{for } x < 2$$
$$p(2) = 1/36$$
$$p(3) = (1/36) + (2/36) = 3/36$$
$$p(4) = (1/36) + (2/36) + (3/36) = 6/36$$
$$p(12) = (1/36) + (2/36) + (3/36) + \ldots + (1/36) = 36/36 = 1$$
$$p(x) = 1 \quad \text{for } x > 12 \qquad \text{(A.2.2)}$$

This may be summarized by:

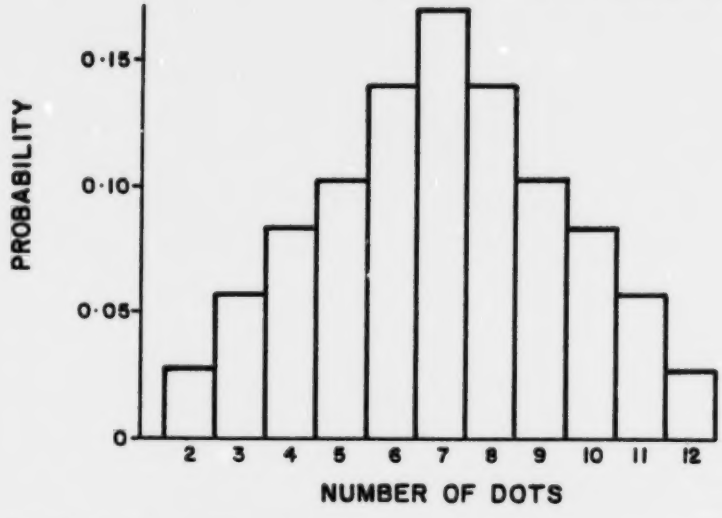
$$p(x) = 0 \quad \text{for } x < 2$$
$$p(x) = x(x-1)/72 \quad \text{for } x = 2, 3, \ldots, 7$$
$$p(x) = (x-4)(21-x)/72 \quad \text{for } x = 8, 9, \ldots, 12$$
$$p(x) = 1 \quad \text{for } x > 12 \qquad \text{(A.2.3)}$$

The graph of this cumulative probability function is shown in Fig. A.2.1(b).

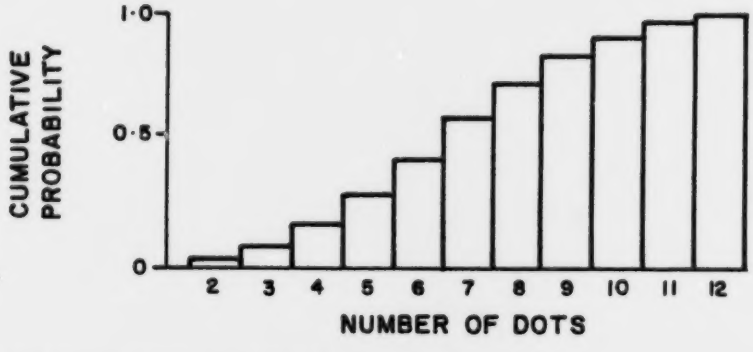
Sometimes the cumulative probability function states the probability that the random variable will be equal to or greater than (rather than equal to or less than) the stated values. This function for the two dice is plotted in Fig. A.2.1(c), and is known as the complementary cumulative probability distribution.

FIGURE A.2.1

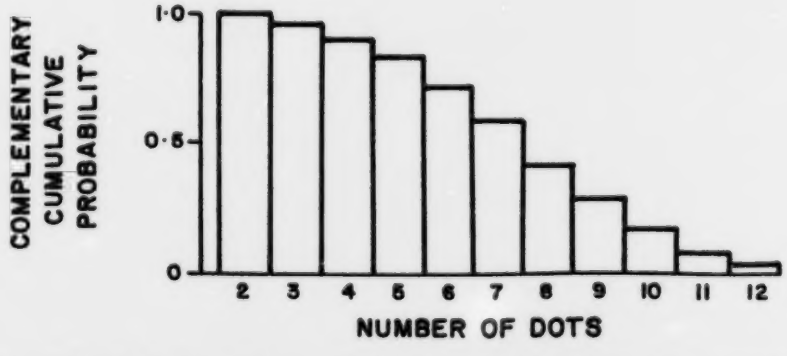
PROBABILITY DISTRIBUTIONS FOR NUMBER OF DOTS ON ROLLING TWO DICE



A) PROBABILITY FUNCTION



B) CUMULATIVE PROBABILITY FUNCTION



C) COMPLEMENTARY CUMULATIVE PROBABILITY FUNCTION

A.3 Continuous Random Variables

Sometimes it is convenient to make the assumption that the random variable may take on the value of any real number over a range. In these cases it is not possible to discuss the probability that the random variable will take on some particular value expressed to an infinite number of decimal places; it is only possible to state the probability of a value within some range. This type of random variable is called continuous, as distinct from the discrete random variables which describe outcomes which may be counted, as described previously with dice, etc.

The probability density function may be used to describe the probability distribution of this kind of random variable. It is a function f(x) such that the probability that the random variable will be between a and b is given by the integral $\int_a^b f(x)dx$ which is illustrated in Fig. A.3.1. The probability is the area under the graph of the probability density function between a and b. It should be noted that this implies that the total area under the probability density function is 1.

The cumulative probability function as described previously may also be used to describe continuous random variables. It is represented as:

$$F(x) = \int_{-\infty}^{x} f(x)\, dx \qquad (A.3.1)$$

and is shown in Fig. A.3.2(b). The value of the function $F(x_1)$ is equal to the area under the probability density function to the left of x_1 as shown in Fig. A.3.2(a). On Fig. A.3.2(b) are also shown the lower quartile, the median, and the upper quartile, which are the values of the variable x for which F(x) is .25, .5 and .75 respectively. These values are also sometimes called the 25th, 50th and 75th percentiles. For completeness the complementary cumulative probability function is included as Fig. A.3.2(c).

FIGURE A.3.1

A PROBABILITY DENSITY FUNCTION

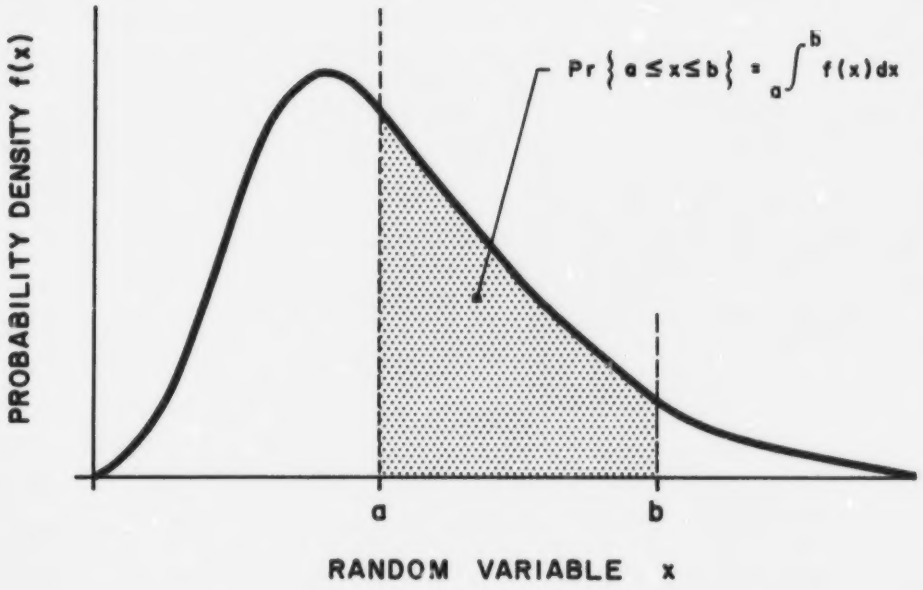
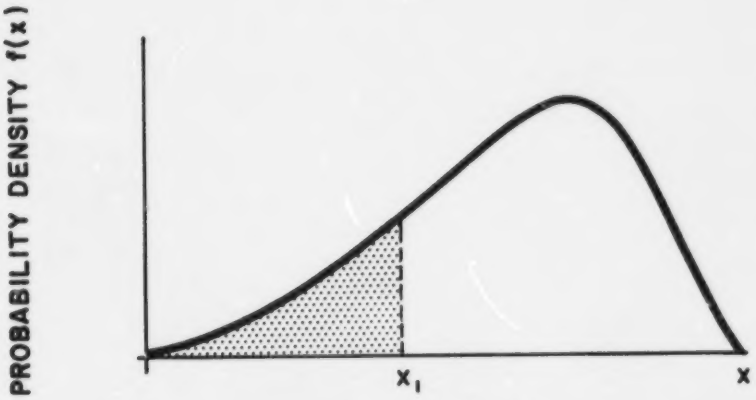
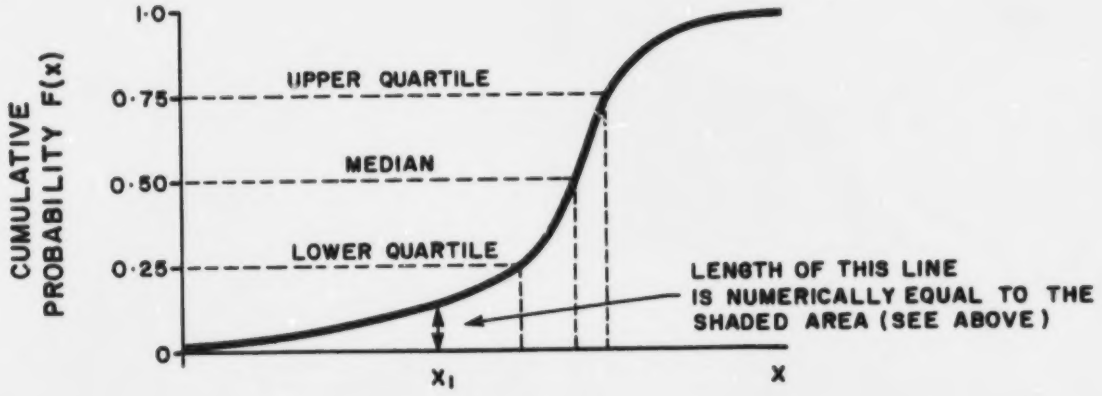


FIGURE A.3.2

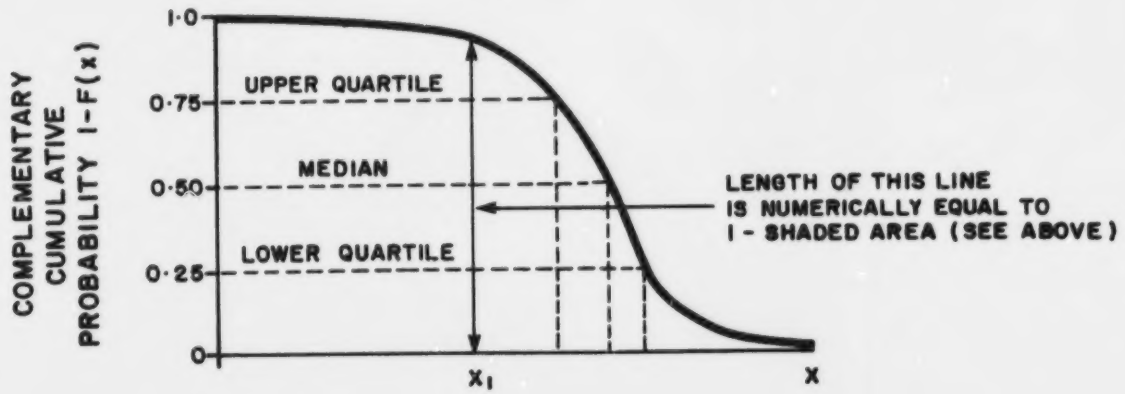
RELATIONSHIPS BETWEEN PROBABILITY DENSITY, CUMULATIVE PROBABILITY AND COMPLEMENTARY CUMULATIVE PROBABILITY FUNCTIONS



A) PROBABILITY DENSITY FUNCTION



B) CUMULATIVE PROBABILITY FUNCTION



C) COMPLEMENTARY CUMULATIVE PROBABILITY FUNCTION

A.4 Examples of Probability Distributions

The following distributions all have considerable use in the type of work described in this report. They are described below and are sketched in Figures A.4.1 to A.4.6.

UNIFORM

The uniform distribution is a simple continuous random variable for which the probability density is constant over an interval (a,b) and is zero elsewhere. The probability density function is:

$$f(x) = 1/(b-a) \quad \text{for} \quad a \leq x \leq b$$
$$f(x) = 0 \quad \text{otherwise} \qquad \text{(A.4.1)}$$

It is completely specified by giving the maximum and minimum.

LOG UNIFORM

The log uniform distribution has a probability density function for which the probability density of the natural logarithm of the variate is constant over the interval (a,b). Its properties are similar to those of the uniform distribution.

$$f(\ln x) = \frac{1}{(\ln b - \ln a)} \quad \text{for} \quad a \leq x \leq b$$

and

$$f(\ln x) = 0 \quad \text{otherwise} \qquad \text{(A.4.2)}$$

NORMAL

The normal distribution is the most commonly used continuous probability density function. The normal probability density function (pdf) has the form:

FIGURE A.4.1

THE UNIFORM DISTRIBUTION

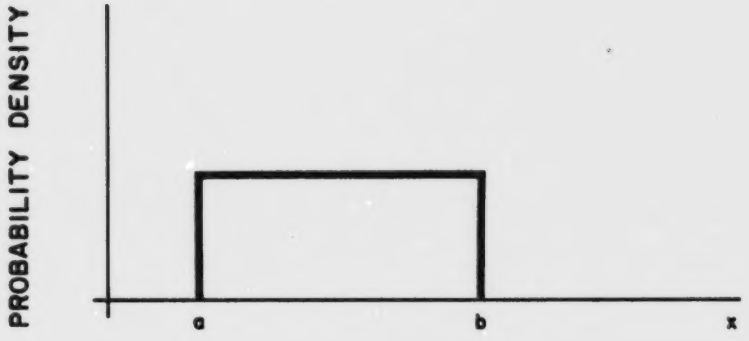


FIGURE A.4.2

THE LOG UNIFORM DISTRIBUTION

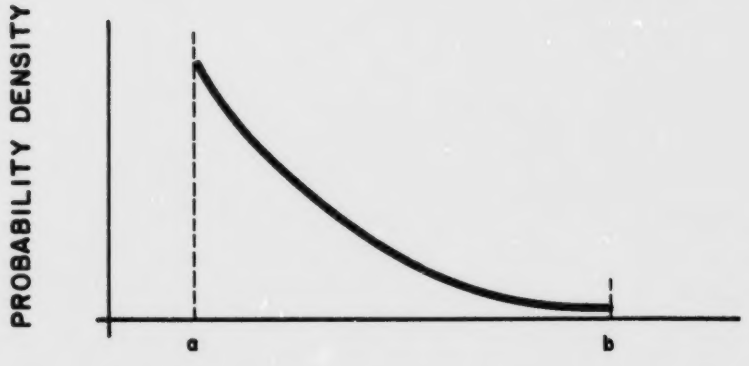


FIGURE A.4.3

THE NORMAL DISTRIBUTION

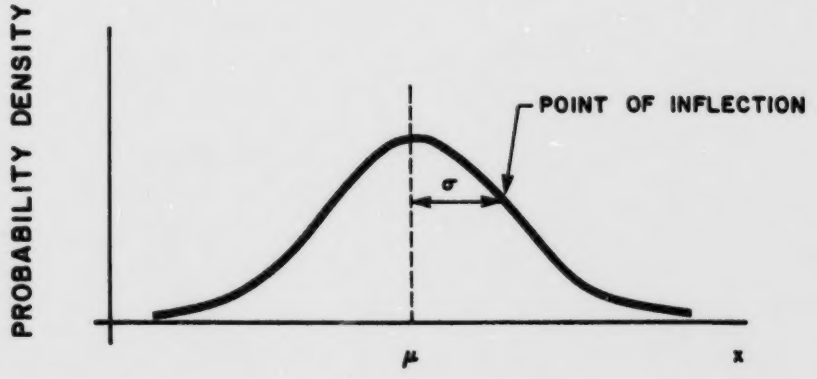


FIGURE A.4.4

THE LOG NORMAL DISTRIBUTION

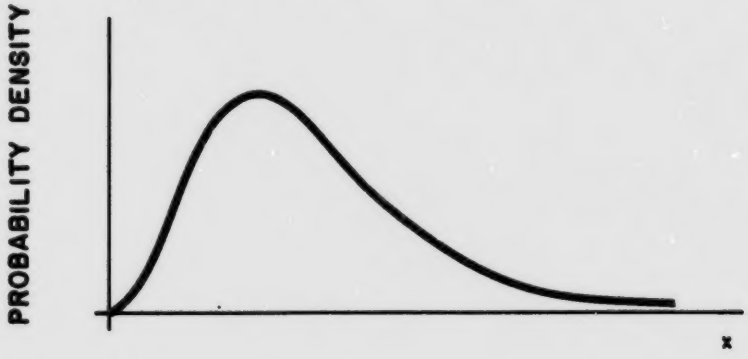


FIGURE A.4.5

THE TRIANGULAR DISTRIBUTION

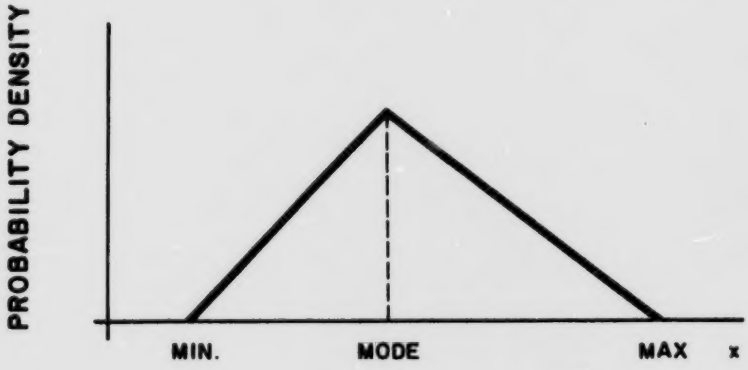
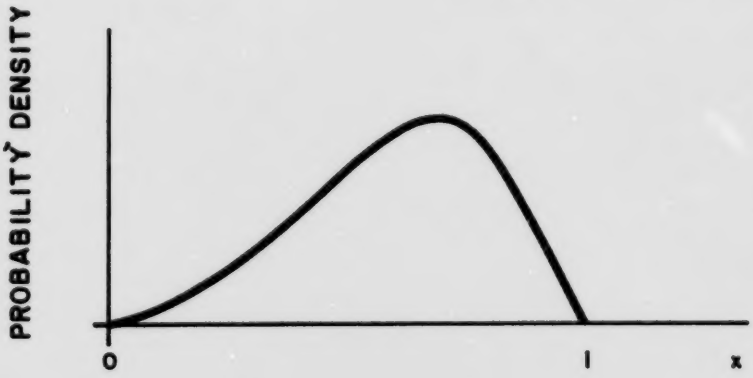


FIGURE A.4.6

THE BETA DISTRIBUTION



$$f(x) = [1/(\sigma\sqrt{2\pi})] \exp[-(x-\mu)^2/2\sigma^2] \text{ for } -\infty < x < \infty \quad (A.4.3)$$
$$-\infty < \mu < \infty$$
$$\sigma > 0$$

where μ and σ are the mean and standard deviation which will be explained in Section A.5.

LOGNORMAL

The lognormal distribution is a continuous distribution that describes a random variable whose logarithm is normally distributed. This distribution is expressed as:

$$f(x) = [1/(x\sigma\sqrt{2\pi})] \exp[-(\ln x-\mu)^2/2\sigma^2] \text{ for } x>0, \sigma>0$$
$$f(x) = 0 \quad \text{otherwise} \quad (A.4.4)$$

where μ and σ are constants which will be explained in Section A.5.

TRIANGULAR DISTRIBUTIONS

The triangular distribution is fully specified by stating the minimum (a) the maximum (c) and the (single) mode (b). The mode may coincide with the maximum or minimum. A mode is a value of the random variable at which the probability is a local or global maximum.

The distribution is specified as:

$$f(x) = \frac{2(x-a)}{(c-a)(b-a)} \quad \text{for } a \le x \le b$$

$$f(x) = \frac{2(c-x)}{(c-a)(c-b)} \quad \text{for } b \le x \le c$$

$$f(x) = 0 \quad \text{otherwise} \quad (A.4.5)$$

BETA DISTRIBUTIONS

The family of distributions called beta distributions is frequently encountered for bounded variates and is described by the following probability density function:

$$f(x) = \left[\frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)}\right] x^{a-1}(1-x)^{b-1} \quad \text{for } 0<x<1,\ a,b>0 \qquad \text{(A.4.6)}$$

and

$$f(x) = 0 \quad \text{otherwise}$$

where $\Gamma(n)$ is the gamma function of n. If n is a positive integer, $\Gamma(n) = (n-1)!$ (i.e. factorial n). The exact form of the distribution varies widely with the values of a and b. If a is less than 1 there is a mode at x=0. If b is less than 1 there is a mode at x=1. If a and b are both less than 1 the curve is U shaped with a mode at each end. Otherwise there is a unique mode which is given by the following:

$$\text{Mode } (x) = \frac{a-1}{a+b-2} \qquad \text{(A.4.7)}$$

The principal advantages of the beta distribution are the wide variety of shapes which can be accommodated by changing a and b, the fact that the distribution is fully specified by stating only a, b, the maximum and the minimum and the fact that the distribution is bounded and thus one can control the occurrence of unrealistic values.

A.5 Expectations and Moments

The expectation of a random variable X is defined as:

$$E(X) = \int_x x\, f(x)\, dx \quad \text{continuous random variable}$$

$$E(X) = \sum_x x\, p(x) \quad \text{discrete random variable} \qquad \text{(A.5.1)}$$

The notation of x under the integral or summation sign shows that the operation is to be done over all possible values of X. Thus E(X) is the average value of X upon infinite observation. The expectation of X is equivalently called the mean, which implies arithmetic mean. The geometric mean is different, and is defined near the end of this section.

The expectation of any mathematical function g(X) of the random variable X is defined as:

$$E[g(X)] = \int_X g(x)\, f(x)\, dx \qquad \text{continuous random variable} \qquad (A.5.2)$$

$$E[g(X)] = \sum_X g(x)\, p(x) \qquad \text{discrete random variable}$$

It may be interpreted as the long term average value of the specified function of the random variable.

By the mathematical analogy to the moments of physics, the kth moment of the random variable X is defined as:

$$\mu_k = E(X^k) \qquad (A.5.3)$$

Some special moments are:

1. The first moment of X is the (arithmetic) mean or expectation of X. By the physics analogy it is the value of X at which the graph of the probability function or probability density function would balance on a knife-edge held perpendicular to the X axis. It is often denoted as μ.

2. The second moment, centred about the mean, is called the variance.

$$Var(X) = E\{[X-E(X)]^2\} = E(X^2) - [E(X)]^2 \qquad (A.5.4)$$

By the physics analogy, the variance is akin to the moment of inerta about an axis through the mean.

The above two forms are mathematically equivalent. The square root of the variance is called the standard deviation and is often denoted as σ. By this notation the variance is σ^2.

The mean μ and the variance σ^2 are very important 'parameters' of the distribution of any random variable. A 'parameter' in statistical usage is a constant which quantifies some property of the system. The mean μ, besides being the long term average value of the random variable, is a location parameter; i.e. it establishes broadly where on a number scale the values of the random variable are to be found.

The variance σ^2 and standard deviation, σ, are scale parameters; they establish the scale or width of the distribution. The variance is the more fundamental quantity but its square root, the standard deviation, has the practical advantage that its units are the same as those of the random variable.

In the examples of Probability Distributions, Section A.4, μ and σ refer throughout to the mean and standard deviation. They are most important in connection with the normal distribution and its modification, the lognormal. Either of these is fully specified by stating μ and σ. The statement:

$$X: N(\mu, \sigma^2) \qquad (A.5.5)$$

means that the random variable X is normally distributed with mean μ and variance σ^2.

Areas under the normal probability density function cannot be found by a simple formula and so tables or the equivalent are required. A convenient standardization is available which greatly simplifies the application of these tables, it is:

$$Z = (X - \mu)/\sigma \qquad (A.5.6)$$

where $X:N(\mu,\sigma^2)$. Then Z is a new random variable which always has the distribution:

$$Z:N(0,1) \qquad \text{(A.5.7)}$$

Some quantitative properties of the normal distribution are:

Interval in Z	Interval in X	% Prob. of Random Variable Being in Interval
- 1 to 1	$\mu - \sigma$ to $\mu + \sigma$	68.27
-1.96 to 1.96	$\mu - 1.96\sigma$ to $\mu + 1.96\sigma$	95.00
-2 to 2	$\mu - 2\sigma$ to $\mu + 2\sigma$	95.45
-2.576 to 2.576	$\mu - 2.576\sigma$ to $\mu + 2.576\sigma$	99.00
-3 to 3	$\mu - 3\sigma$ to $\mu + 3\sigma$	99.73

Like the normal, the lognormal distribution is characterized by μ and σ (see Eqn (A.4.4) and Fig. A.4.4). In fact, if X is lognormally distributed with parameters μ and σ, then:

$$\ln(X):N(\mu,\sigma^2) \qquad \text{(A.5.8)}$$

That is, the natural logarithm of X is normally distributed with mean μ and variance σ^2; hence, the "geometric mean" of X is $\exp(\mu)$, (i.e. e raised to the power μ), and the "geometric standard deviation" is $\exp(\sigma)$.

A.6 Samples

Up to this point random variables have been considered in abstraction, with little reference to reality. In the real world the true distribution of a random variable is only known in trivial situations and a fundamental problem is that of deducing information about it. The basic technique is sampling.

A sample of size n on a random variable X is a set of n independent

observations on X. Two important statistics derived from a sample are the sample mean $\bar{x}$ and the sample variance s^2.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad \text{(A.6.2)}$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{x}) = \frac{1}{n-1} \left[\sum_{i=1}^{n} X_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} X_i \right)^2 \right] \qquad \text{(A.6.3)}$$

where the sample consists of the observations X_1, X_2, ...X_n. They are called statistics because they are mathematical functions of a sample which may be used as estimates, respectively, of the mean μ and variance σ^2 of the random variable.

It is of the utmost importance at this stage to distinguish between, on the one hand, the parameters μ and σ^2 of the distribution of the random variable, which are often called the population parameters, and on the other hand the sample statistics $\bar{x}$ and s^2. The population parameters are fixed for any particular random variable X, but in practical situations are unknown. The sample statistics are easily calculated but vary from one sample to the next. They are not equal therefore to the population parameters but they are commonly used as estimates of them.

Because they vary from one sample to the next, the sample statistics are themselves random variables. They have the following properties:

$$\begin{aligned} E(\bar{x}) &= \mu \\ \text{Var}(\bar{x}) &= \sigma^2/n \\ E(s^2) &= \sigma^2 \end{aligned} \qquad \text{(A.6.4)}$$

That is, on repeated sampling the long-term average values of the sample mean and variance are the population mean and variance, of which they are estimates. Estimates which have this property are called unbiased.

The above expression for Var($\bar{x}$) discloses the important fact that, the larger the sample, the smaller is the variance of the sample mean; i.e. the larger the sample, the narrower is the distribution of the sample mean about the population mean. Hence the sample mean becomes a better estimate of the population mean as the sample size becomes larger.

The distribution of the sample mean has a further and most important property: as the sample size becomes larger the <u>form</u> of the distribution of the sample mean approaches the normal distribution. This statement is called the Central Limit Theorem and it establishes the basis for the importance of the normal distribution. The theorem is subject to mild mathematical conditions which are of no importance in practice. The above information about the sample mean may be summarized by the statement.

$$\bar{x}: N(\mu, \sigma^2/n) \text{ in the limit as } n \to \infty \qquad \text{(A.6.5)}$$

With finite n, the above statement is a good approximation in favourable cases for n as small as 4 or 5. In unfavourable cases it is still usually approximately true if n is greater than 25 or 30.

If the distribution of $\bar{x}$ is standardized according to equation (A.5.6) we get

$$Z = (\bar{x}-\mu)/(\sigma/\sqrt{n}): N(0,1) \qquad \text{(A.6.6)}$$

This expression is useful for confidence intervals and significance tests. The corresponding transformation using s, the <u>sample</u> standard deviation in place of σ is:

$$t = (\bar{x}-\mu)/(s/\sqrt{n}) \qquad \text{(A.6.7)}$$

This is also useful for confidence intervals and significance tests. Since, however, s is a random variable while σ is a constant, the distribution of t, on repeated sampling, is wider than the normal. This distribution is called

the Student's t-distribution after its discoverer W. S. Gossett, who published under the pseudonym "Student". The plot of the probability density function is a bell curve which looks much like the normal.

A.7 Frequency Diagrams

The preceding section dealt with the relations of the population parameters μ and σ with the sample statistics $\bar{x}$ and s. For some purposes it is valuable to use the information of a sample to find out something about the probability distribution. This can be done by the use of sample frequency diagrams, often called histograms, and of cumulative frequency diagrams. Fig. A.7.1(a) shows the sample histogram and Fig. A.7.1(b) the cumulative frequency diagram for an actual experiment in which two dice were rolled 50 times. The complementary cumulative frequency diagram is shown in Fig. A.7.1(c). Notice that these resemble Figure A.2.1, which shows the corresponding probability functions for the random variable which counts the number of dots on rolling two dice. The sample frequency curves are not identical to the population probability curves but they would be expected to resemble the population ones more closely with larger and larger samples.

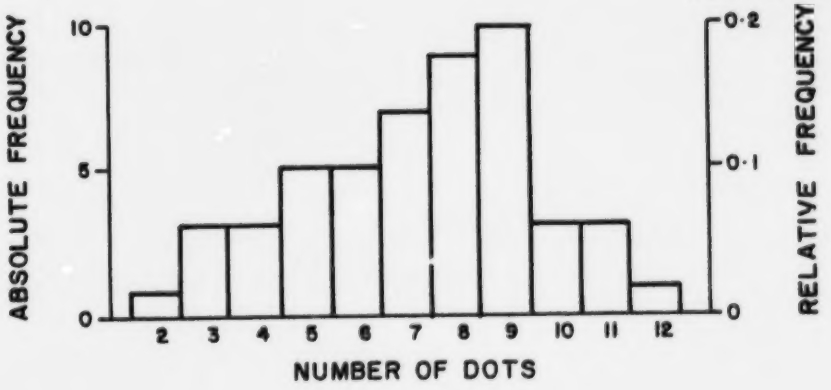
A.8 Point Estimates and Interval Estimates

As described previously, $\bar{x}$ and s^2 may be used as estimates of μ and σ^2, but because of sampling variation they are not, in general, exactly correct. Estimates like this which simply state a single value are called point estimates and have the serious defect that they give no information about their accuracy.

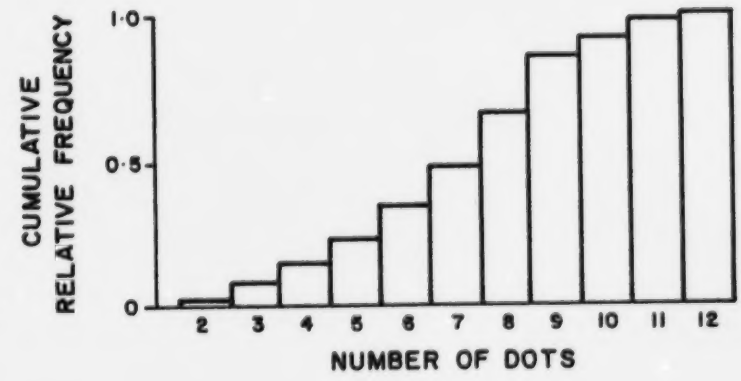
A better way to estimate the value of a parameter is through the use of interval estimates, which in some manner place upper and lower bounds on the value of the parameters. Confidence intervals form one type of interval estimate. They define an interval within which the value of the estimated parameter is believed, with a stated percent confidence, to lie. The confidence level is usually high, at least 90%.

FIGURE A.7.I

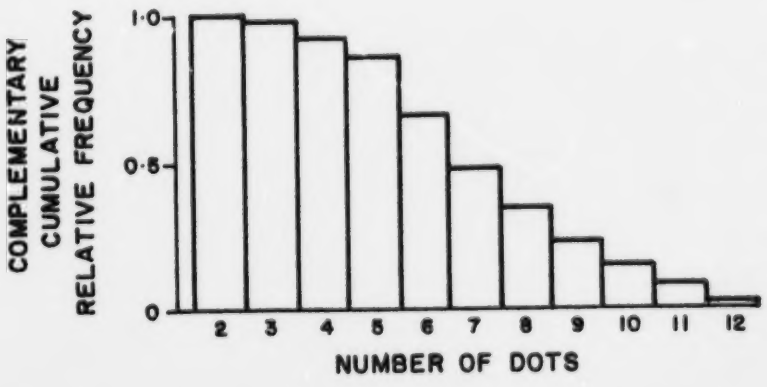
EXPERIMENTAL FREQUENCY DISTRIBUTIONS
FOR NUMBER OF DOTS ON ROLLING TWO DICE
(50 ROLLS)



A) SAMPLE FREQUENCY



B) CUMULATIVE RELATIVE FREQUENCY



C) COMPLEMENTARY CUMULATIVE RELATIVE FREQUENCY

The word "confidence" in this context is a very carefully defined technical term the meaning of which will be developed shortly.

If a sample of size n has been observed on a random variable X, which has known variance σ^2, then a confidence interval at some specified confidence level a for the mean μ may be determined as:

$$\bar{x} - w_a \sigma / \sqrt{n} \leq \mu \leq \bar{x} + w_a \sigma / \sqrt{n} \qquad (A.8.1)$$

or by

$$\bar{x} \pm w_a \sigma / \sqrt{n} \qquad (A.8.2)$$

The first describes the confidence interval while the second states the confidence limits. If σ^2 is unknown the confidence interval (with some increase in average width) is:

$$\bar{x} - t_{a,n-1}\, s/\sqrt{n} \leq \mu \leq \bar{x} + t_{a,n-1}\, s/\sqrt{n} \qquad (A.8.3)$$

or

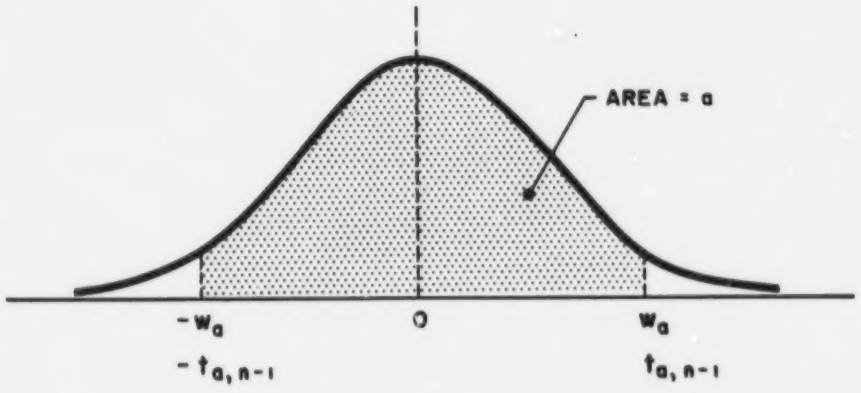
$$\bar{x} \pm t_{a,n-1}\, s/\sqrt{n} \qquad (A.8.4)$$

In the above, w_a and $t_{a,n-1}$ are the values of a standardized variate having respectively the normal distribution (with n degrees of freedom) and the Student t-distribution (with n-1 degrees of freedom). In both cases they are the values of a variate symmetrically placed which enclose an area a between them as indicated on Fig. A.8.1.

The meaning of a confidence interval at confidence level a is a most important consideration. A confidence interval describes the unknown but constant value of a parameter in terms of an interval which is calculated from the observations of a sample. On repeatedly taking the same sized sample under the same conditions, and calculating confidence limits each time, the confidence limits will change from sample to sample while the true value of

FIGURE A.8.1

DEFINITION OF w_α AND $t_{\alpha,n-1}$



NOTE:

FOR w_α THE CURVE IS THE PROBABILITY DENSITY FUNCTION FOR THE STANDARD NORMAL DISTRIBUTION i.e. A RANDOM VARIABLE DISTRIBUTED AS N(0,1)

FOR $t_{\alpha,n-1}$ THE CURVE IS THE PROBABILITY DENSITY FUNCTION FOR THE STUDENT t- DISTRIBUTION WITH n-1 DEGREES OF FREEDOM

the parameter will remain constant. However, if the number of samples is indefinitely large, the fraction of confidence limits which enclose the true value approaches a. The actual stated confidence interval is always calculated from a single sample, but its explanation is in terms of the imagined scheme of indefinitely repeated sampling.

An important property of the confidence intervals given by Equations (A.8.1) to (A.8.4) is that while their derivation depends on X being normally distributed, they are in fact very robust with respect to departure from normality. This means in practice that as long as the sample size is say 5 or more, little attention need be paid to the distribution of X. This property is a consequence of the Central Limit Theorem.

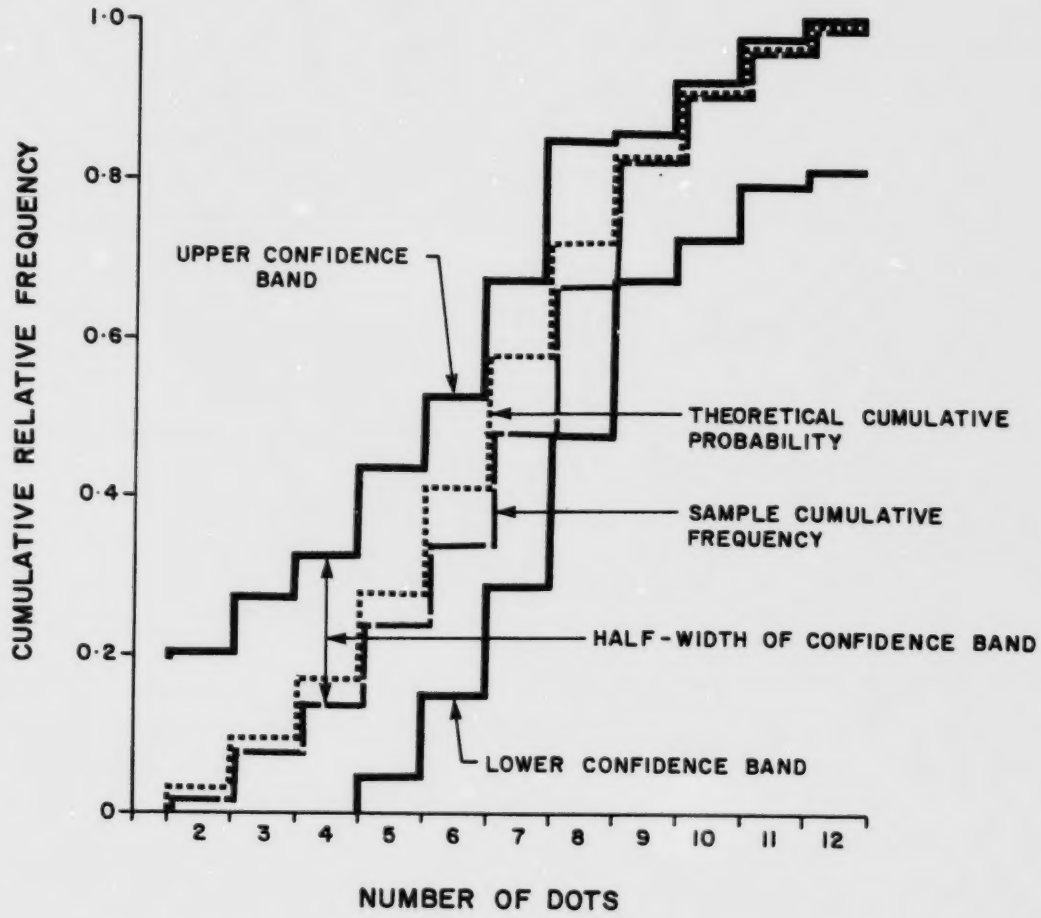
Another type of confidence interval may be used to indicate the true cumulative probability function when the cumulative frequency diagram from a sample has been drawn. Probably the most useful technique for this is that of Kolmogorov and Smirnov (see Crow, Davis and Maxfield, 1960). It allows a staircase-like figure to be drawn above and below the cumulative frequency plot (see Fig. A.8.2) which establishes a confidence region for the cumulative probability function with basically the same meaning as the confidence intervals already described. That is, if the sampling were to be repeated indefinitely under the same conditions, the region between the staircase figures, which varies from sample to sample, would completely enclose the true cumulative probability function, which remains stationary, for a fraction of the samples which is numerically equal to the confidence level. Fig. A.8.2 shows the 95% confidence interval for the cumulative probability function for the number of dots on rolling two dice. The data are the same as those used for Fig. A.7.1. As Fig. A.8.2 indicates, the 95% confidence intervals do, in this case, completely enclose the true cumulative probability distribution.

A.9 A Significance Test for Comparing Sample Means

Significance tests are very frequently used in statistics. A simple one for

FIGURE A.8.2

95% CONFIDENCE BOUNDS FOR A CUMULATIVE PROBABILITY FUNCTION (NUMBER OF DOTS ON ROLLING TWO DICE)



comparing two sample means will be described. Suppose X_1, and X_2 are normally distributed random variables with a common known variance σ^2 and with means μ_1, and μ_2 respectively. It is desired to test if their means are the same.

The first step is to set up the null hypothesis (H_o) that the two means are the same. The word null implies that we assume the proposition we are testing to be true; the word hypothesis here implies that we make the assumption purely for testing purposes. Thus, under the null hypothesis X_1 and X_2 both have the same mean, say u. Then, given H_o:

$$X_1: N(\mu, \sigma^2) \qquad \text{(A.9.1)}$$

$$X_2: N(\mu, \sigma^2)$$

It follows from this that if we have a sample of size n_1 on X_1 and size n_2 on X_2, then

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} : N(0,1) \qquad \text{(A.9.2)}$$

where $\bar{x}_1$ is the sample mean of X_1 calculated by Equation (A.6.2). Similarly, for $\bar{x}_2$.

We now adopt the rule that we reject H_o if $|Z| > w_a$ (see Fig. A.8.1 for definition of w_a). Since large values of $|Z|$ are associated with large differences in the sample means, there must be some value of $|Z|$ beyond which we cannot believe that the two population means are the same (i.e. H_o). The rule establishes such a value with the property that we will only wrongly reject H_o with probability (1-a). When we reject H_o we say, at confidence level a, that there is a significant difference between the two population means μ_1, and μ_2.

Note that if $|Z| \leq w_a$ we cannot say we have proven μ_1 and μ_2 to be equal.

We can only say that we have failed to prove, at confidence level a, that they are different. A null hypothesis, as a simple statement, can never be proven to be true by statistical means.

Often a similar test is needed where σ^2 is unknown. The test criterion then is:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} \qquad \text{(A.9.3)}$$

This differs only from (A.9.2) in that the estimate s^2 is used in place of σ^2. It is calculated from:

$$s^2 = \frac{\sum_{i=1}^{n_1} (X_{1i}-\bar{x}_1)^2 + \sum_{i=1}^{n_2} (X_{2i}-\bar{x}_2)^2}{n_1 + n_2 - 2} \qquad \text{(A.9.4)}$$

where X_{1i}, i=1, 2, ...n, are the observations on X_1, etc. The rule for rejecting H_o is that it is rejected when $|t| > t_{a,n_1+n_2-2}$ where t_{a,n_1+n_2-2} is the symmetrically placed value of the Student's t-distribution on $(n_1 + n_2-2)$ degrees of freedom which encloses a probability a as shown on Fig. A.8.1. This rule has the same confidence property as does the rule for σ^2 known.

With or without σ^2 known, this test, like confidence intervals already described for means, while based on the assumption that X_1 and X_2 are normally distributed, is remarkably robust with respect to departure from normality. This means that as long as the sample sizes are reasonable, say greater than 5 or 6, the form of the distribution of X_1 and X_2 is of little concern.

Of more concern is the equality of the variance of X_1 and X_2. If they cannot be assumed equal, the test is called the Behrens-Fisher test and is more complicated.

A.10 Subjective Probability

Up to this point, this exposition has dealt with random variables as if they were the result of some physical process like rolling dice. Information about them has been in terms of probability which, as applied, may be called objective probability because, at least in principle, any probabilistic statement may be verified by performing the physical experiment to which the probability statement applies and counting up the various types of outcome.

There is another kind of probability which is used in sentences like, "Poor John has only one chance in 10 of ever walking again." There is only one John and it is not possible even in principle to evaluate the probability by repeated observations on him. Such a statement, if used responsibly, means that the speaker feels as strongly about John walking as he does about a favourable outcome in some physical operation which has a 10% objective probability of success. Thus he has communicated his strength of belief in John's walking in terms of a probability. This kind of probability is called subjective probability. It cannot be verified, like objective probability, by counting up outcomes in connection with the proposition being considered. It is verified by comparing degrees of belief in the proposition with degree of belief in the outcome of a physical operation which has a known probability of success.

Thus subjective probability provides a means of quantifying a person's opinions. Note that the worth of the opinion is neither increased nor decreased by the use of subjective probability to describe it. Subjective probability is just as capable of communicating an erroneous opinion as it is a correct one. Once the correctness of the opinion has been accepted, however, a statement of subjective probability quantifies the degree of belief which the person to whom the statement is addressed is entitled to hold.

The mathematics of probability is an abstraction which applies equally well to objective and subjective probability. Consequently all of the material on

random variables, their probability distributions, moments, expectations, variances, etc., have as valid an interpretation in terms of subjective probability as they have in objective probability. Of course, sampling theory does not apply directly to the theory of subjective probability because repetitive independent observations on the same random variable are not possible.

A.11 Probabilistic Analysis

This is a technique which involves a mathematical model of some physical system. The model is a mathematical expression to which information is given in the form of numerical values of one or more input variables and it produces numerical values of one or more output variables. In this report systems are described in which the input variables are such quantities as wind velocity, radium concentration in air, and the like, while output variables might be radiation dosages to various receptors, etc. The mathematical model is the set of mathematical formulas which describes the output variables in terms of the input ones.

Most of the input variables are not known with certainty. The uncertainty arises because of normal experimental error in their measurement or because sometimes actual measurements are impossible and judgement must be used. Also, some input variables such as wind velocities are inherently variable with time. Consequently it is impossible for most input variables to provide a single number which adequately describes its value at a specified time.

As described in Section A.10, however, it is relatively easy to describe the full state of knowledge about any input variable in the form of a distribution of subjective probability. Ordinarily one of the forms described in Section A.4 and shown in Figs. A.4.1 to A.4.6 is used, but it is quite possible to accommodate distributions of any arbitrary form.

Having obtained the input distributions it is necessary to find the corresponding values of the output variables. Necessarily there will be

uncertainty in the input variables. It can be assessed by taking advantage of the mathematical equivalence of objective and subjective probability through the use of a statistical sampling, or Monte Carlo, procedure.

The Monte Carlo procedure is a repetitive one. At each repetition, a single observation is made on each input variable according to its probability distribution. These values of the input variables are substituted into the mathematical model to obtain a single value for each output variable. Then when several repetitions have been made, the set of numerical values for each output variable may be considered a sample on an objective random variable corresponding to that output variable. From the sample, the mean, variance, or probability function may be estimated as desired. Having obtained this information, it may be interpreted as describing the output variable in terms of subjective probability, telling us in quantitative terms what we are entitled to believe about the output variable. Fig. A.11.1 provides an illustration of this process. The OPANAL computer program, described in Appendix B, carries out the Monte Carlo calculations and (usually) estimates the probability density function for the output variables.

The number of Monte Carlo repetitions required for a particular problem depends on what information is being obtained on the output variables. If only the mean (probably the best overall one-shot value) is desired, 100 repetitions (often called trials) is more than ample, while at the other extreme, 500 is the absolute minimum for useful information if the full probability density function is required. As many as 10,000 trials are used in some circumstances. Of course, the more trials, the more precisely the output information is determined.

It is important to distinguish between inherent uncertainty in the value of the output variable which stems from uncertainty in the input variables, and uncertainty in the information about the output variables which is the result of sampling error caused by the Monte Carlo process. Suppose we are using OPANAL to determine the probability density function of a particular output variable. If we perform only 100 trials, the frequency diagram, which is an

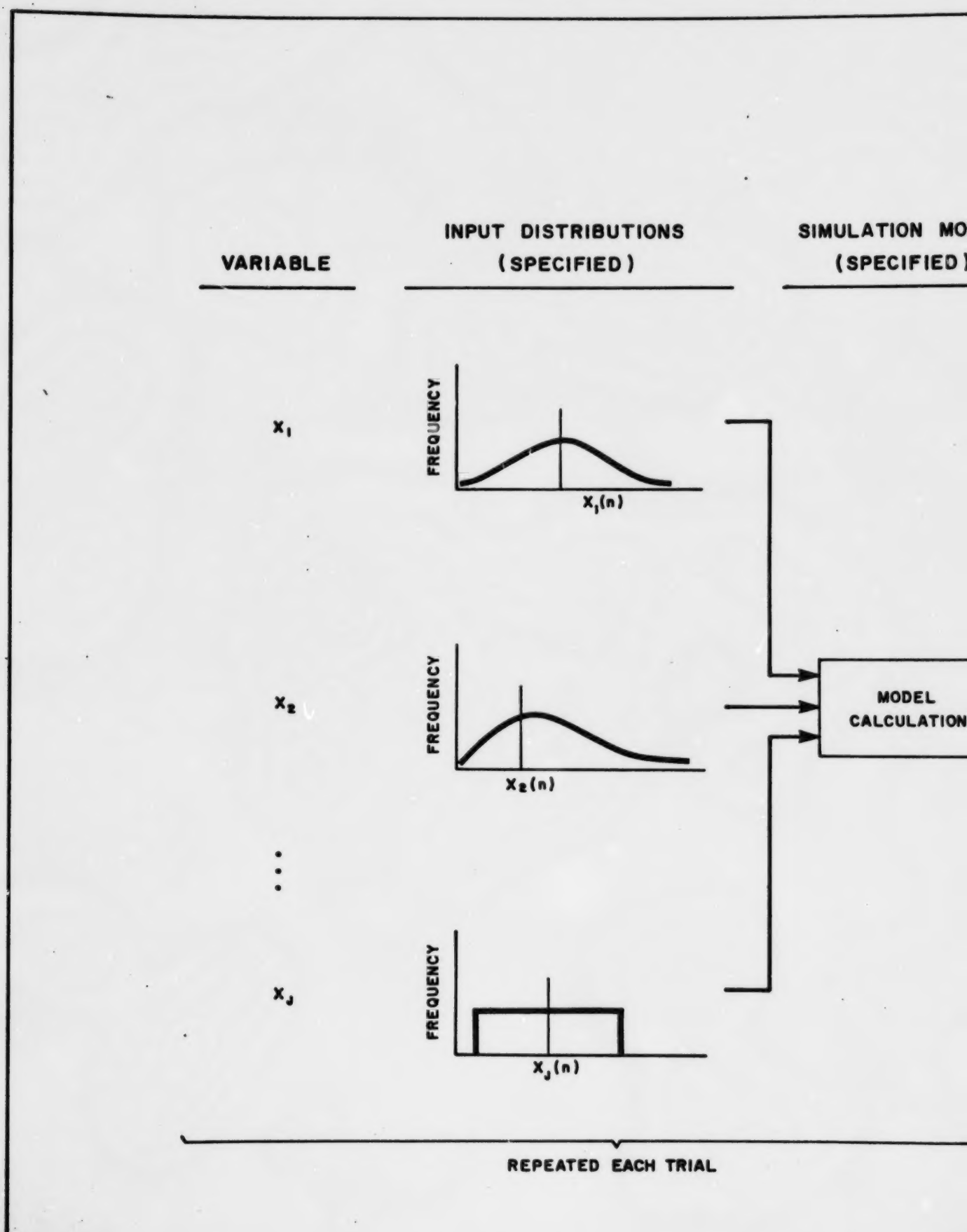

VARIABLE
INPUT DISTRIBUTIONS
(SPECIFIED)
SIMULATION MO
(SPECIFIED)
X_1
FREQUENCY
$X_1(n)$
X_2
FREQUENCY
$X_2(n)$
X_J
FREQUENCY
$X_J(n)$
MODEL
CALCULATION
REPEATED EACH TRIAL

FIGURE A.II.1

PROBABILISTIC ANALYSIS

EL

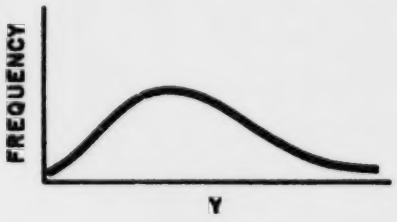
OUTPUT FREQUENCY
SINGLE TRIAL

OUTPUT DISTRIBUTIONS
FROM (n) REPEATED TRIALS
(UNKNOWN)

$Y_1 (X_1, X_2, \ldots X_J)$

$Y_n (X_1, X_2, \ldots X_J)$

$Y_n (X_1, X_2, \ldots X_J)$



SINGLE VALUE
FROM n^{th} TRIAL

SAMPLING DISTRIBUTION OF
OUTPUT VALUES OF SIZE n

estimate of the probability density function, will be very ragged and it will be very difficult to estimate from it the probability density function. If 10,000 trials are used, the frequency diagram will be smooth and will represent the probability density function closely. However, the actual value of the output variable which will be attained in the real world is still uncertain and the uncertainty is communicated quantitatively by the probability density function. This is the inherent uncertainty in the value of the variable. When only 100 trials were used there was an additional uncertainty in the determination of the probability density function.

In a similar vein, if only the mean of an output variable is being reported, it is appropriate to present a confidence interval. This confidence interval has exactly the same meaning as described in Section A.8 and the uncertainty described by the confidence interval is only sampling error resulting from the Monte Carlo process. It bears no relation to the inherent uncertainty in the variable which would be conveyed by its probability density function. For example, the confidence interval for the mean would be very narrow if a very large number of Monte Carlo trials had been used, even if the distribution of the random variable was broad.

Conversely, let us say a very large number of trials had been used and the probability density function was therefore known very closely. Suppose we are interested in the 2.5 and 97.5 percentiles of this variable. It would be inappropriate to call them "95% confidence limits." They are not, because they represent inherent uncertainty in the variable, not sampling error. They could accurately be called 95% probability limits.

A.12 A Practical Example

Thus far, in this appendix, the statistical concepts have been presented with little or no reference to their applications in this study. To illustrate how some of the concepts might be used to present and interpret the results of a probabilistic analysis, an example of actual probabilistic results is provided here.

Figure A.12.1 shows relative frequency and cumulative relative frequency distributions for the estimated annual dose to the receptor (see Section 4.2). The relative frequency distribution A.12.1(a) is shown as a histogram and the cumulative relative frequency distributions A.12.1(b) and (c) as smoothed curves. (Properly, the cumulative relative frequency would also be a step function however, the frequency distribution A.12.1(a) would become smoother as the number of trials and bins increased from the 100 trials and 20 bins used to produce the results shown in Figure A.12.1(a)).

The frequency distribution shown on Figure A.12.1(a) is the result of 100 trials using the Monte Carlo random sampling process. The distribution appears somewhat skewed with the mode shifted to the left (low dose) and a long tail extending to the right (higher dose). Various sample statistics are shown on the figure. As always in this type of work the output distribution shown in Figure A.12.1 may be interpreted in terms of subjective probability, which in effect represents quantitatively what one is entitled to believe about the output variable, dose.

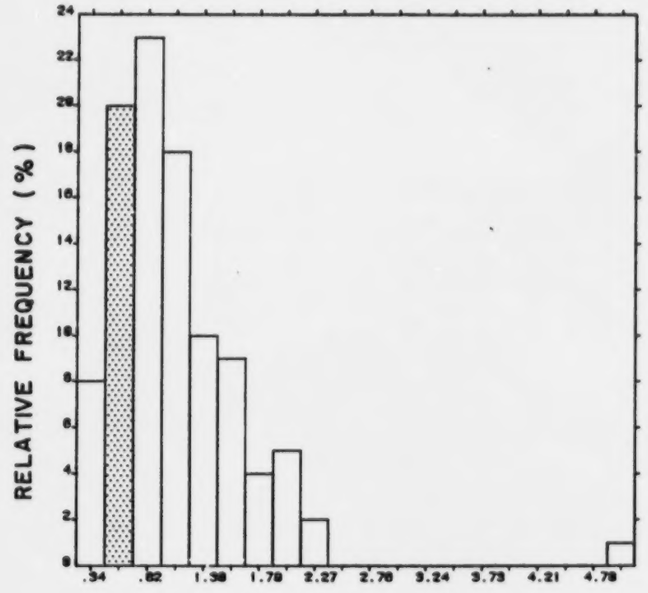
The maximum dose obtained during the 100 random trials of 4.74×10^{-4} $Sv.a^{-1}$ is about 4.6 times larger than the (arithmetic) mean dose of 1.03×10^{-4} $Sv.a^{-1}$. The minimum dose obtained from the hundred trials of 3.29×10^{-5} $Sv.a^{1}$ is about 1/3 of the (arithmetic) mean dose.

The subjective probability interpretation of this is that one is entitled to believe that the maximum dose, observed from 100 trials, of 4.74×10^{-4} $Sv.a^{-1}$ is unlikely to be exceeded more than 1 time in a hundred. The best point estimate of dose (the mean) is 1.03×10^{-4} $Sv.a^{-1}$. It must be realized that with only 100 random trials there is considerable sampling error associated with the figures. The sampling error can be reduced, if desired, by obtaining larger samples.

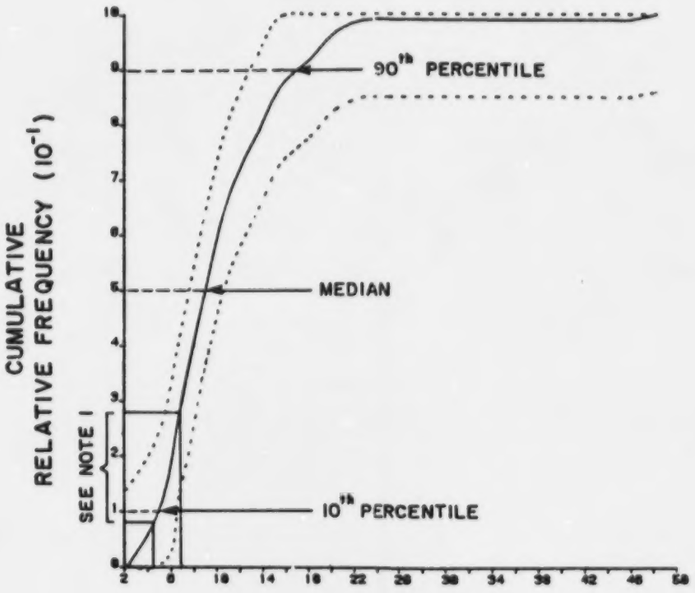
The 10th, 50th and 90th percentiles are illustrated on the cumulative curve, Figure A.12.1(b) (and analogous 90th, 50th and 10th percentiles on the

FIGURE A.12.1

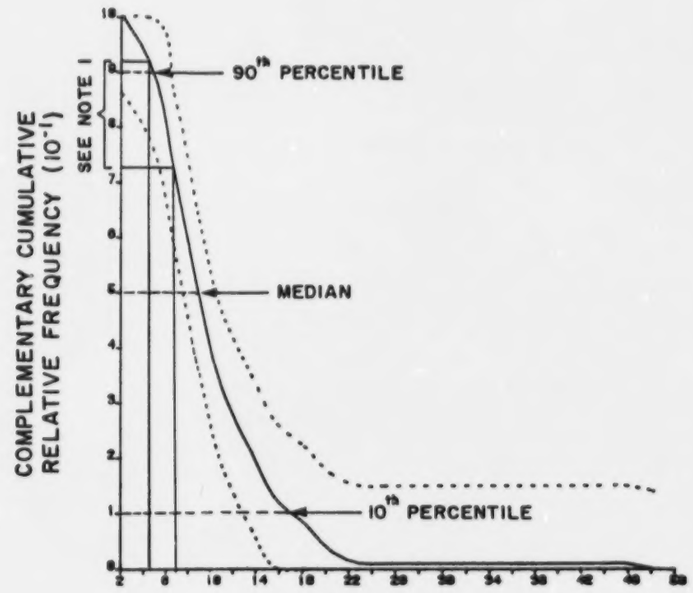
DISTRIBUTION OF DOSE TO RECEPTOR



A) RELATIVE FREQUENCY



B) CUMULATIVE RELATIVE FREQUENCY



C) COMPLEMENTARY CUMULATIVE RELATIVE FREQUENCY

NOTE:

1. THE LENGTH OF THIS LINE IS NUMERICALLY EQUAL TO SHADED AREA (SEE FIGURE A)

2. MINIMUM = 3.29×10^{-6}
   MAXIMUM = 4.74×10^{-4}
   ARITHMETIC MEAN = 1.03×10^{-4}
   ARITHMETIC S.D. = 5.80×10^{-5}
   GEOMETRIC MEAN = 9.16×10^{-5}
   GEOMETRIC S.D. = 1.60

complementary cumulative curve Fig. A.12.1 (c)). From the graph it can be seen that the probability of the dose being in the quite reasonable range (subject to sampling error as described above) of about 5.0×10^{-5} Sv.a^{-1} to about 16.5×10^{-5} Sv.a^{-1} is 80%. Again this may be interpreted to mean that 80% of the trials gave dosage values lying in that range. Alternatively it may mean that one is entitled to feel 80% strongly that the dose under the circumstances being considered would be in the stated range. Similarly we can feel only 10% strongly that the dose will be above 16.5×10^{-5} Sv.a^{-1}.

For illustrative purposes the observed probability of obtaining a dose in the range of 0.46×10^{-4} to 0.7×10^{-4} Sv.a^{-1} is shown as the shaded area on Figures A.12.1(a) and as equivalent line lengths on Figures A.12.1 (b) and (c).

The dashed lines on Figures A.12.1(b) and (c) represent 95% confidence bounds on the respective cumulative distribution plots. Strictly speaking, the confidence intervals are discrete staircase curves (as shown in Figure A.8.2). However, the upper and lower confidence bounds can be calculated at any point and hence, they are shown as smooth curves for convenience. This confidence interval implies that we are 95% confident that the true (unknown) cumulative distribution will be completely enclosed by the dashed lines.

References (Appendix A)

Crow, E.L., Davis, F.A., and Maxfield, M.W., 1960. "Statistics Manual." Dover Publications, New York.

Lindley, D.V., 1965. "Introduction to Probability and Statistics From a Bayesian Viewpoint Vol. I". Cambridge University Press.

Walpole, R.E. and Myers, R.H., 1978. "Probability and Statistics for Engineers and Scientists (2nd Edition)." MacMillan Publishing Co. Inc.

APPENDIX B

OPANAL: A BRIEF DESCRIPTION

APPENDIX B - OPANAL: A BRIEF DESCRIPTION

The OPANAL computer program, one of the several available probabilistic assessment tools, was selected for implementation in this preliminary study. OPANAL was originally developed by Dr. P. Reilly at the University of Waterloo and was modified slightly for the present application. The program consists of a set of routines which generate random variables, and do the bookkeeping for probabilistic analysis. It is written as a subroutine itself with variable dimensions to accommodate any number of input and ouput parameters.

OPANAL contains subroutines to randomly sample from six parametric distributions, namely: normal, lognormal, uniform, log-uniform, triangular and beta. As well, it has the capacity to deal with completely arbitrary, user-specified input distributions. These are specified as a series of x and y coordinates.

A starter routine must be supplied, which calls the OPANAL main routine and reads certain parameters from an input data file. The user is also required to provide a subroutine which performs the calculations specific to the particular application, along with the appropriate input data. The main OPANAL routine consists of four phases:

Phase 1: -reads information from the input data file (number of trials in phases 2 and 3, parameter titles, details of arbitrary and beta input variates).
-calculates normalized cumulative probability functions for the user specified input variates.
-initializes variables as required

Phase 2: -runs a preliminary set of trials (eg. 10-50).
-uses these preliminary results to calculate scaling factors for output histograms.
-writes results to a user specified output file.

Phase 3: -the main functional part of the program.
-runs a large number of trials (e.g. many evaluations of the user specified model).
-writes results of each trial to a user specified post-processor file (distinct from the output file).
-increments cell counts for each variate for which an output histogram is being constructed.

Phase 4: -output histograms are plotted.
-other pertinent summary information (eg. cell counts, scaling factors) is provided for each output variate.
-the raw results from phase 3 remain in a file available for post-processing.

A logic and information flow diagram for the OPANAL program is given in Figure B.1. A variable dictionary, defining some of the parameters referred to in Figure B.1 is also provided (Table B.1).

FIGURE B·1

OPANAL: GENERALIZED LOGIC AND INFORMATION FLOW DIAGRAM

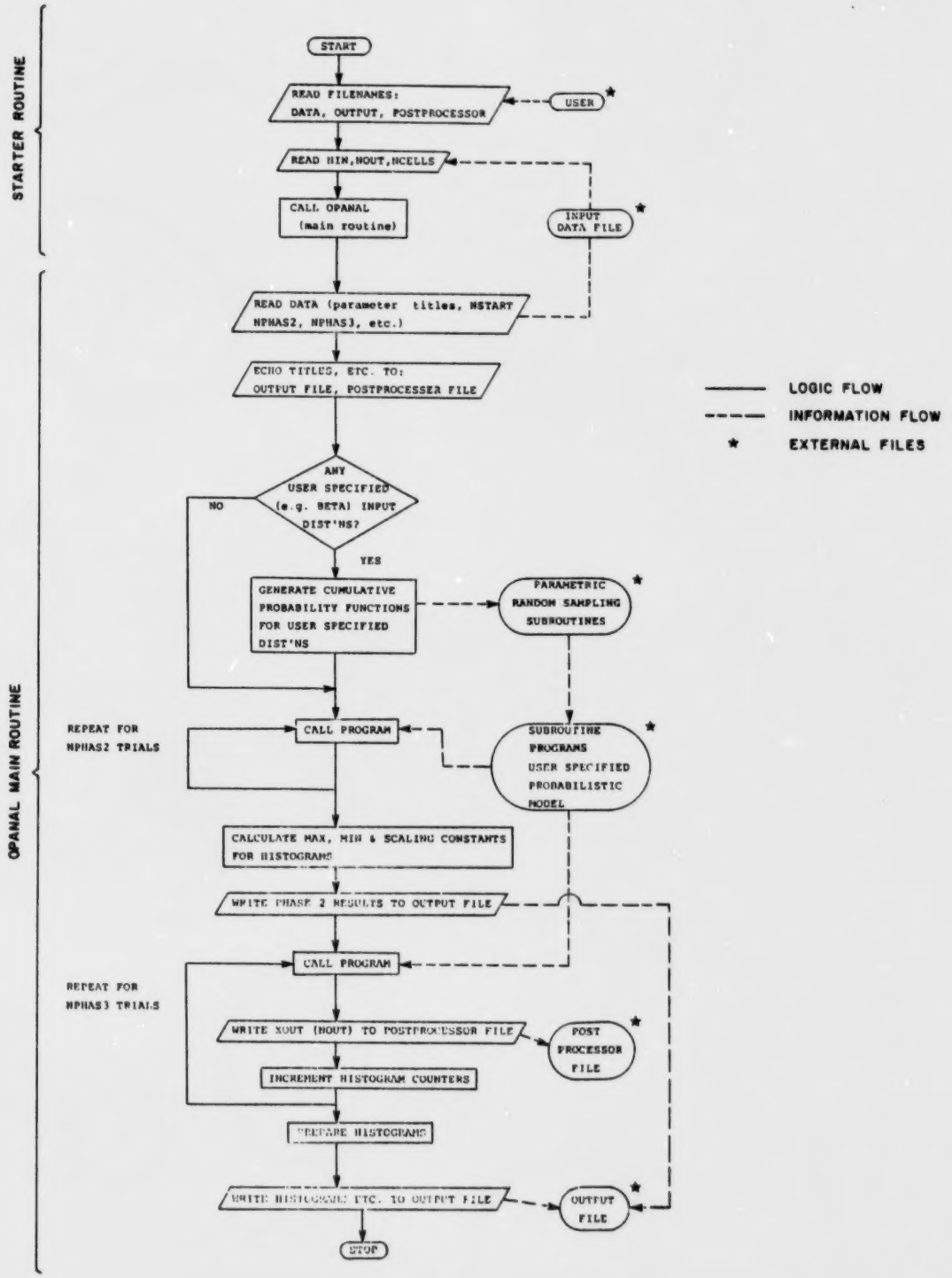


Table B.1

OPANAL Variable Dictionary

Variable Name	Type*	Explanation
CLMAX	V	Maximum values of output variates in Phase 2 and multiplicative scale constants in Phase 3
CPF	M	Row number is number in sequence of variates handled by value subprogram. Column number is cell number, CPF is cumulative probability function for corresponding value of variate
CRAMIN	V	Minimum values of output variates in Phase 2 and additive scale constants in Phase 3
I1	S	General index variables
I2	S	General index variables
ICELL1	S	Index variables for cell numbers
ICELL2	S	Index variables for cell numbers
IDIST1	S	Index variables for distributions
IDIST2	S	Index variables for distributions
NCELLS	S	Number of cells in input and output probability density functions
NCOUNT	M	Rows are output variates. Column number is cell number. NCOUNT is number of counts for corresponding value of output variate
NIN	S	Number of input variates
NPHAS2	S	Number of trials in Phase 2
NPHAS3	S	Number of trials in Phase 3
NOUT	S	Number of output variates
NSTART	S	Random number seed. Odd integer less than 2147483648
V	V	'Scratch-pad' vector
XOUT	V	Output variates
Z1	S	'Scratch-pad' real variables
Z2	S	'Scratch-pad' real variables

*Type V = vector
M = matrix
S = scaler

APPENDIX C

COMPONENT MODELS FOR RADON-222 AND RADIUM-226

APPENDIX C: COMPONENT MODELS FOR RADON-222 AND RADIUM-226

C.1 Introduction

To undertake a preliminary probabilistic assessment of the environmental effects of uranium tailings and to demonstrate the utility of the technique, six pathways were selected for investigation. The pathways were identified on the basis of previous tailings management investigations which showed that the most important exposure pathways for radon and radium are:

Source	Environmental Transfer	Dose
1) radon emanation	to atmospheric dispersion	to receptor inhalation
2) dust erosion	to atmospheric dispersion	to receptor inhalation
3) dust erosion	to vegetation uptake	to receptor ingestion
4) dust erosion	to atmospheric dispersion to vegetation uptake to animal produce	to receptor ingestion
5) radium leaching	to surface water and groundwater	to receptor ingestion
6) radium leaching	to surface water and groundwater to aquatic animals	to receptor ingestion

These pathways were illustrated previously on Figure 2.5 and include the following components: the source term (i.e. the tailings area); several environmental compartments; transfer models; and the receptor.

The concept of the pathways analysis presented on Figure 2.5 is that the dose to a receptor via any one of the pathways is calculated as a series of multiplicative steps. The dose received from several pathways is determined by summing the dose estimates from the individual pathways, that is:

$$X_{10} = \sum_{ij} P_{ij} X_i \quad (\text{for } i \neq j)$$

where:

X_{10}= the total dose to the receptor via all pathways

X_i = the contaminant emission rate from the source term (X_o); or the environmental concentration in the atmosphere (X_1), surface water (X_2), groundwater (X_3), forage or crop vegetation (X_5), animal (caribou) produce (X_6) or aquatic (fish) animals (X_7).

P_{ij}= the environmental dispersion terms from source to atmosphere ($P_{0,1}$), source to surface water ($P_{0,2}$) or source to groundwater ($P_{0,3}$); or alternatively the environmental transfer (conversion or distribution) terms from atmosphere to dose ($P_{1,10}$), atmosphere to vegetation ($P_{1,5}$), surface water to dose ($P_{2,10}$), surface water to fish ($P_{2,7}$), groundwater to surface water ($P_{3,2}$) vegetation to animal produce ($P_{5,6}$), vegetation to dose ($P_{5,10}$), animal produce to dose ($P_{6,10}$) or fish to dose ($P_{7,10}$).

Calculation of receptor exposure using the above equation requires the input of approximate values for the source and environmental compartments (i.e. the X_i blocks) and for the transfer terms (i.e. the $P_{i,j}$ lines).

The determination of input values in pathways analysis normally involves the application of mathematical models to simulate the particular environmental components under consideration. The remainder of this appendix is devoted to documentation of the models adopted in this study to characterize each of the environmental and transfer components of the six specific pathways designated for inclusion in the receptor dose calculations. In all cases, the models selected had been applied in one or more previous pathways type analyses of uranium mill tailings.

C.2 Atmospheric Pathways, Component Models for Radium

C.2.1 Source Term Model: $X_{o,air}$ Source to Air ($Bq.s^{-1}$)

$$EMIT = (RESFAC*RHOTAL*RESDEP)*RACNTA*HSUBD*WIND*(AREXTL)^{1/2}$$

where:

EMIT = source emission rate for $X_{o,air}$ ($Bq.s^{-1}$)

RESFAC = resuspension factor for dust from tailings (m^{-1})

RHOTAL = tailings density ($g.m^{-3}$)

RESDEP = effective resuspension depth (m)

RACNTA = radium concentration in tailings particles ($Bq.g^{-1}$)

HSUBD = initial mixing depth (m)

WIND = annual mean wind speed ($m.s^{-1}$)

AREXTL = exposed tailings surface area (m^2)

Detailed Documentation for $X_{o,air}$ (Radium)

The release rate for dust to the atmosphere (termed $X_{o,air}$) from the tailings surface is assumed to be solely a result of the action of wind on this surface. The source strength is directly dependent on the concentration of radium in that part of the tailings that is available for resuspension.

The source strength calculation is based on the resuspension factor approach (Gilbert 1983, Anspaugh 1975). If desired, the resuspension factor can be made time dependent (it is in fact time dependent but rapidly (<20 years), approaches an asymptotic value), and can be adjusted to account for catastrophic events which modify the tailings surface, such as a forest fire. The resuspension factor is defined as

$$\text{RESFAC} = \frac{\chi}{\text{RHOTAL} * \text{RESDEP}}$$

where:

χ = airborne dust concentration immediately above the tailings within the initial mixing depth ($g.m^{-3}$)

RESFAC = resuspension factor for dust from tailings (m^{-1})

RHOTAL = tailings density ($g.m^{-3}$)

RESDEP = effective resuspension depth (m)

The emission rate for radium in dust from the tailings surface is that given by Gilbert (1983):

$$EMIT = \chi * RACNTA * HSUBD * WIND * (AREXTL)^{1/2}$$

where:

EMIT = source emission rate ($Bq.s^{-1}$)
RACNTA = radium concentration in tailings particles ($Bq.g^{-1}$)
HSUBD = initial mixing depth (m)
WIND = annual mean wind speed ($m.s^{-1}$)
AREXTL = area of exposed tailings (m^2)

Combining the two equations leads to:

$$EMIT = RESFAC * RHOTAL * RESDEP * RACNTA * HSUBD * WIND * (AREXTL)^{1/2}$$

This approach to source term modelling is simple but still quite versatile. The frequency and intensity of high wind speeds is probably more important than the annual mean wind speed. However, a large amount of the uncertainty associated with the calculation is accommodated by the choice of distribution for the resuspension factor, RESFAC.

The structure of this model is such that time dependency in the source emission rate can be incorporated if desired. Two factors, AREXTL, the exposed tailings area, and RESFAC, the resuspension factor will vary with time. The exposed area will gradually decrease as vegetation covers the site. It can increase dramatically if certain catastrophic events such as drought, forest fire, or containment failure occur. The resuspension factor, RESFAC will decrease with time, to an asymptote, as the tailings surface weathers. The small particles available for resuspension will gradually become depleted from the surface layer and RESFAC will decrease. Catastrophic events which disturb the surface lyer and expose unweathered tailings will cause an increase in RESFAC. Thus the dust source term model can readily accommodate both time dependency and catastrophic events if so desired.

C.2.2 Transfer Model: $P_{0,1}$ Atmospheric Transport ($s.m^{-3}$)

$$P_{0,1}=2.032*FRACT*(WIND*DIST*SIGZED)^{-1}*EXP(-((HSUBD^2)*((2*SIGZED^2)-1)))$$

where:

$P_{0,1}$ = transfer factor for radium from the source via atmospheric transport ($s.m^{-3}$)

FRACT = fraction of time the wind blows from source to receptor

WIND = wind speed ($m.s^{-1}$)

DIST = downwind distance (m)

SIGZED = vertical dispersion coefficient (m)

HSUBD = initial mixing depth (source height) (m)

Detailed Documentation for $P_{0,1}$ Atmospheric Transport

The diffusion of material in the atmosphere as it is transported downwind from the source to the receptor is generally called dispersion. Because of its simplicity and widespread application (Hanna 1982, Turner 1967, U.S. NRC 1983) the sector averaged Gaussian plume model was chosen to model the dispersion from the tailings site. The equation is:

$$X = 2.032 \sum_{jk} \frac{n_{jk}}{N} \left[x \; u_k \; \sigma_{zj}\right]^{-1} \exp\left[\frac{-H^2}{2\sigma_{zj}^2}\right]$$

where:

X = the atmospheric concentration at the receptor

x = the downwind distance from the source to the receptor.

u_k = the wind speed of the kth class

σ_{zj} = the vertical diffusion parameter at downwind distance z and stability class j.

H = the height of release of the contaminant

k = wind speed class index

j = atmospheric stability class index

n_{jk}= length of time of the joint occurrence of wind speed class k, and stability class j, in a particular wind direction.

N = total length of time of observation

Assuming that the wind observations are reported for 16 sectors, the radium concentration in the atmosphere at the receptor can be calculated by the following simplified equation:

$$RAAIR = 2.032*EMIT*FRACT*(DIST*WIND*SIGZED)^{-1}*\exp\left[-\frac{HSUBD^2}{2*SIGZED^2}\right]$$

where:

EMIT = the time rate of release of the contaminant ($Bq.s^{-1}$)

FRACT = the fraction of time the wind blows in the sector of interest

WIND = is the mean wind speed in that sector ($m.s^{-1}$)

DIST = the distance from source to receptor (m)

SIGZED= the stability class weighted average variance (m)

HSUBD = the pollutant release height, or the initial mixing depth (m)

RAAIR = the radium concentration in air at the receptor ($Bq.m^{-3}$)

and the transfer factor can be found to be:

$$P_{0,1} = \frac{RAAIR}{EMIT}$$

The equations presented so far in this section are applicable to a point source. In the case of a tailings pile the actual source is not a point source but an area of finite size. In order to accommodate this area the concept of a virtual point source is used. A virtual point source is an imaginary point source located a distance upwind, such that the standard deviation of the plume from the imaginary source is matched to the width of the area source at the source. The virtual set back distance is a function of atmopsheric stability. For the reference tailings site it was calculated by assuming the source was 5×10^5 m^2 in area and was roughly circular. Then for each stability class a setback distance was calculated using well known

procedures (Turner 1967). These distances were combined by taking an average, weighted by the percent frequency of occurrence of each stability class. Thus a mean setback distance of 1600 m was calculated. In all calculations involving DIST, a total distance of 3600 m was used, not the actual distance of 2000 m.

C.2.3 Transfer Model: $P_{1,5}$ Atmosphere to Vegetation ($m^3.g^{-1}$)

Airborne materials can be transferred to the edible portions of vegetables consumed by man or to forage consumed by animals via two mechanisms: direct foliar retention and root uptake. The transfer model used considers only the direct foliar retention. Root uptake was not modelled because of the relatively small contribution it makes to the total vegetation concentration.

Four categories of vegetation are considered: lichen (winter forage), summer forage, above ground vegetables and below ground vegetables.

The following equation is used (U.S. NRC 1982):

$$P_{1,5}=VELDEP*FDEPVE*ESUBV*\left[\frac{(1-\exp(-LAMWTH*TEXPDP)}{YDENST*LAMWTH}\right]$$

where:

VELDEP = the deposition velocity ($m.s^{-1}$)

FDEPVE = the fraction of the total deposition retained on the plant surface (0 to 1)

ESUBV = the fraction of the foliar deposition retained on edible portions of the vegetation (0 to 1)

LAMWTH = a decay constant accounting for weathering losses (s^{-1})

TEXPDP = the duration of plant exposure for vegetation type v to atmospheric deposition (s)

YDENST = the yield density for vegetation type v (g (wet wt).m^{-2})

C.2.4 Transfer Model: $P_{5,6}$ Vegetation to Animal Produce ($g.g^{-1}$)

Radionuclides deposited on winter (lichen) and summer (pasture) forage may be eaten by caribou which in turn may be eaten by man. The transfer of radionuclides from vegetation to meat was modelled as follows (U.S. NRC 1981):

$$P_{5,6} = QVEG*FETOPR*FVEGCN$$

where:

QVEG = intake of vegetation type v by animal type a ($g(dry\ wt).a^{-1}$)

FETOPR = the feed-to-product (p) transfer coefficient for animal type a ($Bq.g^{-1}.Bq^{-1}.a$).

FVEGCN = the fraction of vegetation type v consumed by animal a that is local (0 to 1)

C.2.5 Transfer Model: $P_{5,10}$ Vegetation to Dose ($Sv.a^{-1}.Bq^{-1}.g$)

The annual dose due to ingestion of vegetation was calculated using the following transfer model (U.S. NRC 1982):

$$P_{5,10} = USUBV*DCFING*FVEGGR$$

where:

USUBV = the annual human ingestion rate of vegetation type v ($g.a^{-1}$)

DCFING = dose conversion factor for ingestion ($Sv.Bq^{-1}$)

FVEGGR = the fraction of vegetation type v grown at the contaminated source (0 to 1)

No correction was made for a reduction in the concentration of vegetables due to food preparation (such as washing, peeling and boiling).

C.2.6 Transfer Model: $P_{6,10}$ Animal Produce to Dose ($Sv.a^{-1}.Bq^{-1}.a$)

The transfer model for calculating annual dose due to the consumption of animal produce was (U.S. NRC 1982):

$$P_{6,10} = USUBP*DCFING*FSUBP$$

where:

USUBP = the annual human ingestion rate of animal produce type p ($g.a^{-1}$)

DCFING = dose conversion factor for ingestion ($Sv.Bq^{-1}$)

FSUBP = the fraction of animal produce that is contaminated (0 to 1)

C.2.7 Transfer Model: $P_{1,10}$ Atmosphere to Inhalation Dose (Particulate) ($Sv.a^{-1}.Bq^{-1}.m^3$)

The inhalation dose from radium in suspended particulate is calculated from the relationship:

$$P_{1,10} = OCCFAC*DCFINH*INHALE$$

where:

OCCFAC = occupancy factor or fraction of the year spent at that location (0 to 1).

DCFINH = dose conversion factor for the inhalation of radium ($Sv.Bq^{-1}$)

INHALE = annual inhalation rate ($m^3.a^{-1}$)

The dose conversion factor assumes particles with activity median aerodynamic diameters (AMAD) between 0.1 and 1.0 μm (Johnson and Dunford 1983).

C.3 Surface Water Pathways: Component Models for Radium

C.3.1 Source term Model: $X_{o,water}$ Source to Water ($Bq.s^{-1}$)

i) Radium Concentration in Tailings Porewater and Contaminated Runoff

$$RAPOWA = RAGPTA*0.1*\left\{\left[PHI^2 + \frac{4*KSUBSO}{FACTCO^2}\right]^{1/2} - PHI\right\}$$

where:

RAPOWA = radium concentration in the tailings porewater ($Bq.m^{-3}$)

RAGPTA = radium content of gypsum in the tailings ($Bq.mole^{-1}$)

KSUBSO = solubility equilibrum constant for gypsum ($moles^2.m^{-6}$)

FACTCO = activity coefficient:

for PHI>0.125, FACTCO=0.3

for PHI$\leq$0.125.

$$FACTCO = \exp_{10}\left\{2.4*\left(\frac{PHI}{1000}\right) - \left(\frac{4*(PHI/1000)^{1/2}}{1 + 2*(PHI/1000)^{1/2}}\right)\right\}$$

PHI = concentration of sulphates derived from pyrite oxidation (moles/m^3)

$$= 1.033*NPRCP^{-1}*[(SULPRD*RHOTAL*O2CONC*O2DIFF*(1-KPRIME*\ln((1+KPRIME)*KPRIME^{-1}))]^{1/2}$$

where:

NPRCP = net precipitation (i.e. after evapotranspiration losses) on tailings ($m.a^{-1}$)

SULPRD = average yearly sulphate production rate (moles .$(g.a)^{-1}$)

RHOTAL = tailings bulk density ($g.m^{-3}$)

O2CONC = oxygen concentration at the tailings surface ($moles.m^{-3}$)

O2DIFF = oxygen diffusivity through the tailings ($m^2.a^{-1}$)

$$KPRIME = \frac{K_{O2}}{O2CONC}$$

where:

K_{O2} = half saturation constant for oxygen ($moles.m^{-3}$)

ii) Source Term Emission Rate for Radium and Surface Water ($X_{o,water}$)

$$XOSWRA = RAPOWA*NPRCP*(1.0-FNPRCP)*ARBASN*((3.1536x10^7)^{-1})$$

iii) Source Term Emission Rate for Radium in Groundwater ($X_{o,water}$)

$$XOGWRA = = RAPOWA*NPRCP*FNPRCP*ARBASN*((3.1536x10^7)-1)$$

where:

FNPRCP = fraction of precipitation on the tailings drainage basin that infiltrates and enters the deep groundwater flow below the tailings.

ARBASN = area of the tailings basin watershed (m^2)

XOSWRA = source term emission rate for radium in surface water ($Bq.s^{-1}$)

XOGWRA = source term emission rate for radium in groundwater ($Bq.s^{-1}$)

Detailed Documentation for $X_{o,water}$ (Radium)

Development of a preliminary source term model which describes the mobilization of radium from uranium tailings was undertaken specifically for pyritic tailings undergoing oxidation of the pyritic fraction. Much of the readily leachable radium is present in slightly soluble co-precipitates, such as gypsum ($CaSO_4.2H_2O$), that formed either during tailings neutralization with lime, or by secondary mineralization in the tailings basin. From the leaching studies of Silver and Ritcey (1980), it is estimated that 25% to 50% of the total radium present in tailings is contained in the potentially leachable fraction of the tailings solids. Tailings typically contain approximately 5% gypsum (as $CaSO_4$).

In the preliminary source term model, the concentration of radium in the porewater below the tailings surface is calculated as a function of the concentration of (dissolved) gypsum in the same porewater. Since the molar concentration of gypsum will be approximately equal to the molar concentration of calcium, the radium concentration can be expressed as:

$$[Ra^{2+}] = K[Ca^{2+}]$$

where:

[ ]= concentration of dissolved ions (moles.m^{-3})

K = proportionality constant (equal to the concentration of Ra^{2+} in gypsum)

The solubility of gypsum, and hence the concentration of Ca^{2+} is dependent on the solubility product as follows:

$$\frac{KSUBSO}{\left[f_{ca^{2+}}\right]\left[f_{so_4^{2-}}\right]} = \left[SO_4^{2-}\right]\left[C_a^{2+}\right]$$

where:

KSUBSO = solubility equilibrium constant (25.0 moles^2.m^{-6})

$f_{ca^{2+}}$ = activity coefficient (calcium ion)

$f_{so_4^{2-}}$ = activity coefficient (sulphate ion)

The activity coefficients, assumed to be essentially equal, can be estimated from the Davies equation (Stumm and Morgan 1981):

$$\log \text{FACTCO} = -AB^2 \left\{ \frac{(I)^{1/2}}{1 + (I)^{1/2}} - 0.3I \right\}$$

where:

FACTCO = activity coefficient (f)

I = ionic strength of the solution

A = $1.82 \times 10^6 (T)^{-1.5}$ (=dielectric constant of water)

B = ionic charge

At 10^0C (283.15^0K) the value of 'A' in the above equation is 0.5 (dielectric constant = 83.95)

The ionic strength (I) can be calculated from:

$$I = 1/2 \sum_i C_i z_i^2$$

where

C_i = concentration of ith ionic species

z_i = ionic charge of the ith species

Assuming an equal content of divalent cations and anions, with the sulphate ion being the predominant anion, it can be shown that:

$$I = 4*[SO_4^{2-}]$$

Substituting these values into the expression for calculating the activity coefficients (i.e. the Davies equation):

$$FACTCO = \exp_{10} \left\{ 2.4[SO_4^{2-}] - \frac{4[SO_4^{2-}]^{1/2}}{1+2[SO_4^{2-}]^{1/2}} \right\}$$

In theory, the sulphate concentration in the above equation should include sulphate resulting from both pyrite oxidation and gypsum dissolution. Since the solubility of gypsum is dependent on the activity coefficient, the calculation procedure should be iterative. Recognizing that pyrite oxidation is the major source of sulphate, the ionic strength is assumed to be numerically equal to twice the ionic strength resulting from sulphate production by pyrite oxidation. Although this gives an approximate value for the activity coefficient, it avoids the use of cumbersome trial and error methods to calculate the activity coefficient.

It may be assumed that the limestone ($CaCO_3$) content (up to 1%) of pyritic tailings is rapidly depleted due to acid generation resulting from pyrite oxidation above the water table. Thus, the dissolution of gypsum is the major source of calcium during the initial years. On the other hand, pyrite oxidation of tailings contributes significantly to the sulphate concentration in the porewater. That is:

$$[Ca^{2+}] = [Ca^{2+}]_g$$
$$[SO_4^{2-}] = [SO_4^{2-}]_g + [SO_4^{2-}]_p$$

where:

$[C_a^{2+}]_g$ = calcium derived from dissolution of gypsum

$[SO_4^{2-}]_g$ = sulphate derived from dissolution of gypsum

$[SO_4^{2-}]_p$ = sulphate derived from pyrite oxidation (PHI)

In the relationships expressed above:

$$[Ca^{2+}]_g = [SO_4^{2-}]_g$$

hence, substituting into the equation given previously for the solubility product, the following expression is obtained:

$$\frac{KSUBSO}{\left[f_{Ca^{2+}}\right]\left[f_{SO_4^{2-}}\right]} = ([SO_4^{2-}]_p + [SO_4^{2-}]_g)\ [Ca^{2+}]_g$$

or

$$\frac{KSUBSO}{\left[f_{Ca^{2+}}\right]\left[f_{SO_4^{2-}}\right]} = [SO_4^{2-}]_p * [Ca^{2+}]_g + [Ca^{2+}]_g^2$$

solving this equation for $[Ca^{2+}]$ gives:

$$[Ca^{2+}] = 0.5 * \left\{\sqrt{[SO_4^{2-}]_p + \left(4\ KSUBSO \div \left[f_{Ca^{2+}}\right]\left[f_{SO_4^{2-}}\right]\right)} - [SO_4^{2-}]_p\right\}$$

To solve for the activity coefficient (FACTCO), the sulphates derived from pyrite oxidation must be quantified. The derivation of this quantity depends on a wide variety of physical factors as outlined below.

In pyritic tailings it is convenient to distinguish between the oxygenated zone, comprising the tailings in contact with oxygen, and a redox zone, where chemical reactions proceed in the absence of oxygen. The latter region is influenced by the export products of the overlying oxygenated zone.

The overall oxidation rate (in terms of sulphate generation) is governed by the diffusion of oxygen. The stoichiometry of the oxidative process is as follows (Nordstrom 1982):

$$2\ FeS_2 + 7.5\ O_2 + 7\ H_2O \rightarrow 2\ Fe(OH)_3\ (s) + 8\ H^+ + 4\ SO_4^{2-}$$

The differential equation describing the oxygen concentration at some depth "z" in the tailings, neglecting oxygen transport by barometric or hydraulic "pumping", is given by (Halbert et al 1983):

$$\frac{\partial C_{(z,t)}}{\partial t} = O2DIFF \frac{\partial^2 C_{(z,t)}}{\partial z^2} - \frac{7.5}{4} * SULPRD * RHOTAL$$

where:

$C_{(z,t)}$ = oxygen concentration at depth "z" (moles.m^{-3})
O2DIFF = oxygen diffusivity through tailings ($m^2.a^{-1}$)
SULPRD = sulphate production rate(moles.$(g.a)^{-1}$)
RHOTAL = tailings bulk density ($g.m^{-3}$)
z = depth into the tailings from the surface (m)

For any value of oxygen demand (ie. SULPRD*RHOTAL*7.5/4 > 0), the oxygenated zone in the tailings develops to a finite thickness, λ. Two solutions of the differential equation are as follows.

i) for initial conditions $C_{(z,o)} = 0$;

$$\frac{C_{(z,t)}}{C_{(o,t)}} = 1 - \frac{z}{\lambda}^2 - \text{erf}\left[\frac{z}{2*(O2DIFF*t)^{1/2}}\right]$$

ii) for initial conditions $C_{(z,o)} = C_{(o,t)}$ and $0 \leq z \leq \lambda$;

$$\frac{C_{(z,t)}}{C_{(o,t)}} = \left[1 - \frac{z}{\lambda}\right]^2 + \frac{32}{3\pi} \sum_{n=1}^{\infty} \frac{1}{n^3} \sin\left[\frac{n\pi}{2}\left(\frac{z}{\lambda}\right)\right] \exp\left[-\left(\frac{n\pi}{2\lambda}\right)^2 * O2DIFF * t\right]$$

where:

$C_{(o,t)}$ = oxygen concentration at the interface (moles.m^{-3})
λ = steady state thickness of the oxygenated zone (m)
erf = the error function

The thickness, λ, is determined from steady state conditions ($\partial C_{(z,t)}/\partial t = 0$):

$$\lambda = \left[\frac{8}{7.5} \frac{C_{(o,t)} * O2DIFF}{SULPRD * RHOTAL}\right]^{1/2}$$

The pyrite oxidation rate (SULPRD), is dependent on a number of variables

including the pyrite content, the particle surface area, moisture content, temperature, pH, and oxygen concentration. Although all of these variables may be included in a comprehensive model, in the present case only the dependence on temperature, moisture content and oxygen concentration are considered. The temperature dependence is evaluated by selecting a typical monthly soil temperature distribution and calculating the yearly average oxidation rate as follows:

$$SULPRD = \sum_{i=1}^{12} R_{so} \exp\left[\frac{\Delta E_a}{R\ T_o} - \frac{\Delta E_a}{R\ T_i}\right]$$

where:

SULPRD = average yearly sulphate production rate as SO_4^{2-} (moles.g^{-1}.a^{-1})

R_{so} = monthly sulphate production rate as SO_4^{2-} at reference temperature (moles .g^{-1}.$month^{-1}$)

E_a = activation energy of pyrite oxidation (25.6 kcal.(mole.^0K)$^{-1}$)

R = the universal gas constant (1.987 cal.$mole^{-1}$)

T_o = reference temperature (293 ^0K)

T_i = soil temperature in the "i"-th month (^0K)

The effect of moisture content on SULPRD is obtained by appropriate adjustment of the reaction rate under unsaturated conditions.

The effect of oxygen concentration on the reaction rate (SULPRD) can be described by the following equation (Shumate et al 1971):

$$R_s^* = SULPRD^* \left[\frac{C_{(z,t)}}{K_{02} + C_{(z,t)}}\right]$$

where:

K_{02} = half saturation constant for oxygen (moles.m^{-3})

R_s^* = average yearly sulphate production rate adjusted for the effect of oxygen concentration (moles.g^{-1}.a^{-1})

The numerical solution of the model requires an estimate of the diffusivity coefficient of oxygen through the tailings. In the present case, the following expression is used:

$$O2DIFF = D_o \left[\frac{\varepsilon_a}{\varepsilon_p}\right]\left[\varepsilon_p\right]^{1.5}$$

where:

O2DIFF = diffusivity coefficient of oxygen in unsaturated tailings ($m^2.a^{-1}$)

D_o = diffusivity coefficient of oxygen in air ($m^2.a^{-1}$)

ε_a = airspace porosity at given moisture content

ε_p = porosity (total)

Although transient solutions have been developed for various initial conditions, only the steady state solution is considered at the present. It can be shown that the steady state sulphate generation flux derived from the oxygen flux into the tailings is given by:

$$J_s = \left\{\frac{8}{7.5} * SULPRD * RHOTAL * O2CONC * O2DIFF * \left[1 - KPRIME * \ln\left(\frac{1+KPRIME}{KPRIME}\right)\right]\right\}^{1/2}$$

where:

J_s = the steady state sulphate flux ($moles.m^{-2}.a^{-1}$)

KPRIME = $K_{O2}/O2CONC$

The sulphate concentration due to pyrite oxidation is thus:

$$PHI = \frac{J_s}{NPRCP} = [SO_4^{2-}]_p$$

where:

NPRCP = net precipitation (i.e. after evapotranspiration losses) on the tailings basin ($m.a^{-1}$)

PHI = concentration of sulphate in the tailings porewater ($moles.m^{-3}$)

Hence, the expression given previously for calculating $[Ca^{2+}]$ can be solved and the radium concentration in porewater can be determined from:

$$RAPOWA = K*0.5*\left\{\left(PHI^2 - \frac{4*KSUBSO}{FACTO^2}\right)^{1/2} - PHI\right\}$$

K = Ra content of gypsum = RAGPTA

= 6.8×10^4 to 2.7×10^5 $pCi*mole^{-1}$

= 2,500 to 10,000 $Bq*mole^{-1}$

The equation above applies if the dissolution of gypsum containing the radium is congruent, i.e. if the molar (or mass) ratio of radium to calcium in the liquid phase (porewater) and the solid phase are identical. However, observed concentrations of radium range from 60 to 240 $pCi*L^{-1}$ in tailings porewater in the Elliot Lake region (Golders Associates 1981a and 1981b, Cherry et al 1980, Blair et al 1980). These porewaters are essentially saturated with respect to gypsum, the calcium ion concentration being 1×10^{-2} M to 1.5×10^{-2} M. The observed ratios of Ra^{2+} to Ca^{2+} in the porewater were 6,000 $pCi*mole^{-1}$ to 23,000 $pCi*mole^{-1}$. These ratios are in sharp contrast to Ra^{2+} to Ca^{2+} ratios of 68,000 to 270,000 $pCi*mole^{-1}$ in the bulk solid phase (gypsum). It appears that the dissolution process is incongruent, resulting in an enrichment of the gypsum surface with radium. Consequently, it appears that the Ra^{2+} to Ca^{2+} ratio in the porewater is attentuated with respect to the bulk solid phase. The resultant attentuation factor, expressed as the ratio of the radium content in the liquid phase to the solid phase may range from 0.005 to 0.34. For the purposes of the Phase I probabilistic analysis, an attentuation factor of 0.2 was used. This attentuation factor acts as a direct multiplier in the calculation of the radium concentration in porewater. Hence, the expression give above becomes:

$$RAPOWA = 0.2*\ RAGPTA*0.5*\left\{\left(PHI^2 - \frac{4*KSUBSO}{FACTCO^2}\right)^{1/2} - PHI\right\}$$

$$RAPOWA = 0.1*RAGPTA* \left\{ \left(PHI^2 - \frac{4*KSUBSO}{FACTCO^2} \right)^{1/2} - PHI \right\}$$

C.3.2 Transfer Model: $P_{0,2}$ Source to Surface Water ($s.m^{-3}$)

The transfer model for contaminated surface runoff from the tailings drainage basin initially involves mixing with freshwater flow in a river located some 200 m from the reference site. Additional mixing with freshwater occurs as the river flow moves downstream. The river discharge is further diluted in a lake situated some 2 km downstream of the reference site. Radium removal on suspended inorganic solids and sorption onto biomass occurs in the lake with subsequent deposition to the bottom sediments. These model components are discussed in turn below.

1) River System Component

$$RARTOT = \frac{(RAPOWA*ARBASN*NPRCP*(1.0-FNPRCP))+JFLUX+QSUBI*RATI}{(ARBASN*NPRCP*(1.0-FNPRCP))+QSUBI}$$

where:

RARTOT = concentration of radium in the river at its discharge into the lake ($Bq.m^{-3}$).

RAPOWA = radium concentration in the tailings porewater ($Bq.m^{-3}$)

ARBASN = area of the basin the tailings are located in (m^2)

NPRCP = net precipitation on the tailings ($m.a^{-1}$)

QSUBI = flowrate of the freshwater in the river system, at the point of discharge to the lake, excluding the tailings area discharge. For a net watershed area of 13.5×10^6 m^2 QSUBI was calculated from the relationship NPRCP x 13.5E6($m^3.a^{-1}$)

RATI = concentration of Ra-226 in the river system, upstream of the tailings discharge ($Bq.m^{-3}$)

FNPRCP = fraction of net precipitation (NPRCP) that flows to the river via the groundwater route

JFLUX = contaminant flux via the groundwater route to the river (set equal to zero for the time frame considered in this analysis) ($Bq.a^{-1}$)

This model simply weights the inputs to the river system to arrive at a volume weighted concentration at the mouth of the river just upstream of the lake. When the water leaves the river and enters the lake, it is mixed and portioned into various components.

ii) Lake System Component

$$RALTOT = \frac{QSUBR*RARTOT/ARLKBT+JRAWCO}{\frac{QSUBO}{ARLKBT} + \frac{SIGSS*KSUBD*SSL+SIGBIO*KSUBBI*BIOWCL}{1+KSUBD*SSL+KSUBBI*BIOWCL}}$$

where:

RALTOT = total Radium-226 in the lake water column ($Bq.m^{-3}$)

QSUBR = total flow in the river, at the discharge to the lake including the tailings sub-basin discharge. For a total watershed area of 15.0 x 10^6 m^2 QSUBR was calculated as NPRCP x 15E6 in the model ($m^3.a^{-1}$)

RARTOT = total radium concentration in the river at the point of discharge into the lake ($Bq.m^{-3}$)

ARLKBT = total area of the lake bottom (m^2)

QSUBO = total water outflow from the lake. For a total watershed area of 20.0 x 10^6 m^2, QSUBO was calculated as NPRCP x 20E6 in the model ($m^3.a^{-1}$)

SIGSS = sedimentation velocity of suspended solids ($m.a^{-1}$)

KSUBD = sorption coefficient for radium onto suspended solids ($m^3.g^{-1}$)

SIGBIO = sedimentation velocity of biomass ($m.a^{-1}$)

KSUBBI = transfer coefficient for radium to aquatic organisms ($m^3.g^{-1}$)

The suspended solids level in the water column is calculated as follows:

SSL = QSUBR*SSR/(ARLKBT*SIGSS + QSUBO)

where:

SSL = concentration of suspended solids in the water column ($g.m^{-3}$)

SSR = concentration of suspended solids in the river discharge to the lake ($g.m^{-3}$)

The biomass level in the water column (BIOWCL in the model) is calculated as follows:

$$BIOWCL = PROBIO/(SIGBIO + QSUBO/ARLKBT)$$

where:

BIOWCL = concentration of aquatic biota in the lake water column ($g.m^{-3}$)

PROBIO = net biomass productivity ($g.m^{-2}.a^{-1}$)

The radium flux from the bottom sediment into the water column (JRAWCO) is calculated as follows:

$$JRAWCO = \frac{KSUBL*QSUBR*RARTOT*(ZETA-1)}{QSUBO*(1 + KSUBBI*BIOWCL + KSUBD*SSL)}$$

where:

KSUBL = mass transfer coefficient of radium into the water column from the sediments ($m.a^{-1}$)

ZETA = concentration ratio of dissolved radium between the interstitial sediment water and the lake water.

JRAWCO = flux of radium from the sediments into the water column ($Bq.m^{-2}.a^{-1}$)

Finally it is necessary to calculate the transfer coefficient, as follows:

$$P_{0,2} = RALTOT/XOSWRA$$

C.3.3 Transfer Model: $P_{2,7}$ Surface Water to Aquatic Animals ($m^2.g^{-1}$)

A simple transfer coefficient was used to model the concentration in edible

portions of locally available aquatic animals as follows (U.S. NRC 1982):

$$P_{2,7} = BFRA$$

where:

BFRA = water to edible portion of aquatic animal type f transfer coefficient ($Bq.g^{-1}.Bq^{-1}.m^3$)

C.3.4 Transfer Model: $P_{7,10}$ Aquatic Animal to Dose ($Sv.a^{-1}.Bq^{-1}.g$)

The annual dose received by an individual as a result of eating aquatic animals was modelled using the following parameters (U.S. NRC 1982).

$$P_{7,10} = USUBA*DCFING*FSUBA$$

where:

USUBA = the annual human ingestion rate of aquatic animal type f ($g.a^{-1}$)

DCFING = dose conversion factors for ingestion ($Sv.Bq^{-1}$)

FSUBA = the fraction of aquatic animal type f that is contaminated (0 to 1)

C.3.5 Transfer Model: $P_{2,10}$ Surface Water to Dose (Ingestion) ($Sv.a^{-1}.Bq^{-1}.m^3$)

The annual dose due to ingestion of surface water was modelled as (U.S. NRC 1982):

$$P_{2,10} = USUBW*DCFING*FWATHM$$

where:

USUBW = the annual human intake of water ($m^3.a^{-1}$)

DCFING = dose conversion factor for ingestion ($Sv.Bq^{-1}$)

FWATHM = the fraction of water consumed by a human that is contaminated (0 to 1)

C.4 Groundwater Pathway: Component Models for Radium

The contribution of the groundwater pathway to the dose to the receptor specified for the Phase I investigation was recognized to be an insignificant source given the characteristics assigned to the reference site. Nevertheless, the groundwater pathway was modelled for the purpose of demonstrating the ability to incorporate time dependency into a probabilistic assessment. While time dependency was not incorporated into any of the other component models and the probabilistic assessment runs were carried out for assumed quasi-steady state conditions, the groundwater models were structured to allow the influence of time on radium transport via this pathway to be evaluated.

C.4.1 Source Term Model: $X_{o,water}$ Source to Groundwater ($Bq.s^{-1}$)

Details of the source term model for predicting the rate of radium release to the surface water and groundwater flows exiting the tailings drainage basin were presented previously in Section C.3.1. To summarize, the model for calculating the rate of release of radium from the tailings area to the groundwater flow component which exits the basin as a deep subsurface flow was given as:

$$XOGWRA = RAPOWA*NPRCP*FNPRCP*ARBASN/(3.1536x10^7)$$

where

XOGWRA = source term release of radium to the groundwater ($Bq.s^{-1}$)
RAPOWA = radium concentration in the tailings porewater ($Bq.m^{-3}$)
NPRCP = net precipitation (after evapotranspiration losses) ($m.a^{-1}$)
FNPRCP = fraction of precipitation on the tailings drainage basin that infiltrates and enters the deep groundwater flow below the tailings.
AR3ASN = area of the tailings basin watershed (m^2)

C.4.2 Transfer Model: $P_{3,2}$ Groundwater to Surface Water ($m^3.m^{-3}$)

The radium flux by the groundwater pathway at some distance from the tailings

basin is given by:

$$JFLUX = GWCONC*QVOLSR$$

where:

JFLUX = radium flux via the groundwater pathway at some distance from the source ($Bq.a^{-1}$)

GWCONC = concentration of radium in the groundwater at some distance from the source ($Bq.m^{-3}$)

QVOLSR = the volumetric rate of groundwater flow from the tailings basin ($m^3.a^{-1}$)

= ARBASN*NPRCP*FNPRCP

The concentration of radium in groundwater, GWCONC, is calculated using a groundwater transport model for simulation of the transport of contaminants from the uranium tailings site to the adjacent river. The component model incorporates simplified one-dimensional steady-state groundwater flow conditions with contaminant-specific retardation parameters to simulate the time dependent transport of contaminants through the groundwater system. The key output of the model is the contaminant flux to the surface water receptor, which, for the reference site, is defined to be a distance of only 200 m from the tailings basin.

In this simplified model, it is assumed that the volumetric rate of porewater seepage from the tailings into the groundwater system is equal to the volumetric rate of groundwater discharge (QVOLSR) to the surface water receptor in the zone influenced by the tailings area. Contaminant concentrations in the groundwater during flow from the tailings to the river are reduced by adsorption on the geologic materials and by radioactive decay. Consequently, the contaminant flux to the surface water (flux = volumetric flow * contaminant concentration) will be reduced by these processes. Large-scale mixing or dilution of the tailings porewater exiting the bottom of the tailings area with fresh groundwater is not accounted for in the model. Although this process would reduce contaminant concentration, the volumetric flow rate would be correspondingly increased so that this process is of no consequence when the contaminant flux rates are calculated.

The groundwater transfer model first calculates the groundwater (DARCY) velocity as:

$$DARCY = HYDCON*HYDGRD$$

where:

DARCY = the Darcy velocity ($m.a^{-1}$)

HYDCON = the hydraulic conductivity ($m.a^{-1}$)

HYDGRD = the hydraulic gradient ($m.m^{-1}$)

Input parameters for the longitudinal dispersion, adsorption and radioactive decay of the contaminants are calculated based on:

$$DISPLT = DISPER*DARCY/POROSG$$

where:

DISPLT = the longitudinal dispersion ($m^2.a^{-1}$)

DISPER = the dispersivity (m)

DARCY = the Darcy velocity ($m.a^{-1}$)

POROSG = the porosity of the geologic material ($m^3.m^{-3}$)

and

$$RETFAC = 1 + (RHOBAS*KSUBDG/POROSG)$$

where:

RETFAC = the retardation factor

RHOBAS = the bulk density of the geologic material along the groundwater flow path ($g.m^{-3}$)

KSUBDG = the distribution coefficient for radium on geologic material along the groundwater flow path ($m^3.g^{-1}$)

The contaminant concentration at a distance (DISTG) from the tailings area is calculated for transient conditions using:

$$GWCONC = \frac{RAPOWA}{2} * EXP\left[\frac{DARCY*DISTG}{2*POROSG*DISPLT}\right] * \Big\{EXP(-DISTG*BETA)*$$

$$ERFC\left[\frac{DISTG-((DARCY/(POROSG*RETFAC))^2+(4*LAMRA*DISPLT/RETFAC))^{1/2}*TIME}{2*(DISPLT*TIME/RETFAC)^{1/2}}\right]$$

$$+EXP(DISTG*BETA)*$$

$$ERFC\left[\frac{DISTG+((DARCY/(POROSG*RETFAC))^2+(4*LAMRA*DISPLT/RETFAC))^{1/2}*TIME}{2*(DISPLT*TIME/RETFAC)^{1/2}}\right]\Big\}$$

where:

RAPOWA = radium concentration in tailings porewater ($Bq.m^{-3}$)

DISTG = the distance from the tailings (m)

LAMRA = the radioactive decay constant for radium (a^{-1})

TIME = time (a)

ERFC = is the complementary error function

and BETA is defined as:

$$BETA = \left[\left[\frac{DARCY}{2*POROSG*DISPLT}\right]^2 + \left[\frac{LAMRA*RETFAC}{DISPLT}\right]\right]^{1/2}$$

C.5 Atmospheric Pathways; Component Models for Radon

C.5.1 Source Term Models for Radon: $X_{o,air}$ Source Emission Rate ($Bq.s^{-1}$)

RNSOUR = ARTLSR*RACNTA*LAMRN*EMANAT*LSUBT*RHOTAL*
((NSUBW*WTSFLX+(12-NSUBW))/12)*exp(-1*(THICK/LSUBC))

where:

RNSOUR = source emission rate of radon, $X_{o,air}$ ($Bq.s^{-1}$)

ARTLSR = total tailings surface area (m^2)

RACNTA = radium concentration in the tailings ($Bq.g^{-1}$)

EMANAT = emanation coefficient

RHOTAL = tailings bulk density ($g.m^{-3}$)
LAMRN = radioactive decay constant for radon (s^{-1})
LSUBT = radon diffusion length in tailings (m)
THICK = clean cover thickness (m)
LSUBC = radon diffusion length in the cover (m)
NSUBW = number of winter months (freezing conditions) per year
WTSFLX = winter-to-summer radon flux ratio (dimensionless)

The radon diffusion length, LSUBT, in tailings may be calculated as follows:

$$LSUBT = (D/(LAMRN*P))^{1/2} \text{ (m)}$$
$$LSUBT = (EBKDFC/LAMRN)^{1/2}$$

where:
P = porosity of tailings
EBKDFC = effective bulk diffusion coefficient for radon ($m^2.s^{-1}$)
= 0.106×10^{-6} exp (-0.261*MPCH2O)
MPCH2O = percent water (by weight) in tailings
D = bulk diffusion coefficient of radon ($m^2.s^{-1}$)

Detailed Documentation for the Radon Source Term, $X_{o,air}$

The following describes a model for the simulation of the rate of release of radon from uranium mill tailings. This release rate is used as the source term in the air dispersion model that calculates radon concentration in air attributable to the tailings at the receptor location. The theory of radon release is well presented in many publications and is not discussed in detail here. For example, the reader may refer to the U.S. NRC's Final Generic Environmental Impact Statement on Uranium Milling (U.S. NRC 1980) for further information or to Chambers et al (1980) for the modelling of radon source terms in the context of Canadian uranium tailings areas.

The radon flux or rate of radon release from the surface of a tailings pile to the atmosphere is dependent on several factors including: the rate at which

radon is produced through the decay of the parent radium; the fraction of radon which escapes the solid matrix in which it was formed and enters the tailings pore space (the emanation coefficient); and the rate at which the radon in the pore space (interstitial radon) migrates to the tailings surface. Radon migration is in turn dependent on both diffusion and transport (ie. bulk flow) phenomena. In the following discussions, the tailings pile is considered to be homogeneous with respect to radium content, water content and particle size, and infinitely deep (>3 to 4 m) in terms of radon diffusion. Although in practice differential settling of tailings particles after discharge to a tailings pile may occur and radium concentration may be a function of particle size, the distinctions are minor relative to the accuracy of the radon model which attempts to predict the overall average (over a year) rate of release of radon from the tailings pile.

i) Radon Production

The radon production rate (q) in a porous radium-bearing material can be expressed as:

$$q = RACNTA*RHOTAL*EMANAT*LAMRN/P$$

where:

q = radon production rate ($Bq.s^{-1}.m^{-3}$ of pore space)
RACNTA = tailings radium concentration ($Bq.g^{-1}$)
RHOTAL = tailings density ($g.m^{-3}$)
EMANAT = emanation coefficient
P = porosity of tailings
LAMRN = radon decay constant ($2.1 \times 10^{-6} s^{-1}$)

The emanation coefficient, EMANAT, varies widely but a typical value of 0.2 has often been used to model radon releases from tailings.

ii) Radon Diffusion

The diffusion of radon through a homogeneous, porous medium can be modelled

using classical diffusion theory. The one-dimensional equation for the radon concentration in a tailings pile is:

$$\frac{\partial C}{\partial t} = \frac{\partial D}{P} * \frac{\partial^2 C}{\partial X^2} - C*LAMRN + B$$

where:

C = radon concentration ($Bq.m^{-3}$ of pore space)
D = bulk diffusion coefficient ($m^2.s^{-1}$)
B = radon source strength ($Bq.s^{-1}.m^{-3}$)
X = distance (m)
P = porosity of tailings
LAMRN = radon decay constant (s^{-1})

The term D/P can be called the effective diffusion coefficient, EBKDFC. Short term variations in the transport of radon (as opposed to pure diffusion) through the tailings that are caused by pressure, windspeed, and temperature fluctuations are often included in the measured values of EBKDFC.

iii) Radon Flux

The steady-state condition ($\partial C/\partial t = 0$) can be solved to give the radon flux J_o at the surface of the tailings:

$$J_o = RACNTA*EMANAT*RHOTAL*LAMRN*LSUBT$$

where:

J_o = radon flux at the tailings surface ($Bq.m^{-2}.s^{-1}$)
LSUBT = radon diffusion length in tailings (m)
= $(EBKDFC/LAMRN)^{1/2}$

and the other terms are as previously defined.

The diffusion of radon in tailings is greatly affected by the moisture content of the tailings. The radon flux from water-saturated tailings can be reduced by a factor of 100 or more relative to the radon flux from dry tailings. A

reasonable empirical correlation between the effective diffusion coefficient and moisture content (Rogers et al 1980) is:

$$EBKDFC = 0.106 \times 10^{-6} \exp(-0.261 * MPCH2O)$$

where:

MPCH2O = percent moisture content (by weight)

More complex empirical relationships between moisture content and diffusion have been derived but their use is not required here to demonstrate the overall probabilistic modelling technique.

iv) Exhalation from Covered Tailings

The flux (J) of radon through a covered tailings pile can be approximated by:

$$J = J_o \exp(THICK/LSUBC)$$

where:

Jo = radon flux at tailings surface ($Bq.m^{-2}.s^{-1}$)

THICK = cover thickness (m)

LSUBC = diffusion length of radon in cover (m)
= $(DC/LAMRN)^{1/2}$

DC = effective diffusion coefficient in cover material ($m^2.s^{-1}$)

and it is assumed that the cover contains no radium. LSUBC for typical soils ranges from 0.5 to 1.5 m.

v) Winter/Summer Correction

The previous discussions have been concerned with radon flux from tailings under normal conditions. However, freezing conditions and snow cover greatly attenuate the radon flux because of the much smaller diffusion coefficient of ice relative to water. The effect is dependent on the moisture content of the tailings (Zettwoog et al 1981). This effect can be accommodated by using a seasonally-weighted average for radon flux and a winter/summer flux ratio.

The annual average flux J_{AVG} is then given by:

$$J_{AVG} = J_s \ [NSUBW*WTSFLX + (12-NSUBW)]/12$$

where:

J_s = summer flux ($Bq.m^{-2}.s^{-1}$)

NSUBW = number of winter months (freezing conditions) per year

WTSFLX = winter/summer flux ratio

= 0.02 nominally (Zettwoog et al 1981)

C.5.2 Transfer Model $P_{0,1}$ Source to Atmosphere ($s.m^{-3}$)

The source-to-atmosphere transfer terms for particulate matter or radon are the same. In developing the transfer model, it was assumed that the materials being dispersed were of neutral buoyancy; generally a good approximation for both radon and the particulate matter released from tailings.

C.5.3 Transfer Model $P_{1,10}$ Atmosphere to Inhalation Dose (Radon) ($Sv.a^{-1}.Bq^{-1}.m^3$)

The inhalation dose from radon is calculated from the relationship:

$$P_{1,10} = OCCFAC*DCFIRN$$

where:

OCCFAC = occupancy factor or fraction of the year spent at that location (0 to 1)

DCFIRN = dose conversion factor for the inhalation of radon (= 7.2×10^{-5} $Sv.a^{-1}.Bq^{-1}.m^3$)

The dose conversion factor assumes a radon daughter/radon equilibrium factor of 0.5, and a dose equivalency of 4.8 WLM and 0.05 Sv.

REFERENCES (Appendix C)

Anspaugh, L.R., Shinn, J.J., Phelps, P.L., and Kennedy, N.C., 1975. "Resuspension and Redistribution of Plutonium in Soils." Health Physics 29, pp. 571-582.

Blair, R.D., Cherry, J.A., Lim, T.P. and Vivyurka, A.J., 1980. "Groundwater Monitoring and Contaminant Occurrence at an Abandoned Tailings Area, Elliot Lake Ontario". Proceedings of the First International Conference on Uranium Mine Waste Disposal, Vancouver, May 19-21.

Chambers, D.B., Lowe, L.M., Cassaday, V.J., Nantel, J.H. and Archibald, J.F., 1980. "Development and Use of Radon Source Terms in Environmental Impact Assessments". Presented at the 25th Annual Meeting of the Health Physics Society, Seattle, Washington.

Cherry, J.A., Blackport, R.J., Dubrovsky, N., Gillham, R.W., Lim, T.P., Murray, D., Reardon, E.J. and Smith, D.J.A., 1980. "Subsurface Hydrology and Geochemical Evolution of Inactive Pyritic Tailings in the Elliot Lake Uranium District, Canada." Proceedings of the Third Symposium on Uranium Tailings Management, Colorado State University, Fort Collins, Colorado, Nov. 24-25.

Gilbert, T.L., Chee, P.C., Knight, M.J., Peterson, J.M., Roberts, C.J., Robinson, J.E., Tsai, S.Y.H., Yuan, Y.C., 1983. "Pathways Analysis and Radiation Dose Estimates for Radioactive Residues at Formerly Utilized MED/AEC Sites". U.S. Department of Energy.

Golder Associates, 1981a. "Hydrogeological Investigation, Williams Lake Management Area." Report to Denison Mines Limited.

Golder Associates, 1981b. "Additional Hydrogeochemical Monitoring, 1981, Williams Lake Management Area." Report to Denison Mines Limited.

Halbert, B.E., Scharer, J.M., Knapp, R.A., and Gorber, D.M., 1983. "Determination of Acid Generation Rates in Pyritic Mine Tailings". Presented at the 56th Annual Conference of the Water Pollution Control Federation, Atlanta, Georgia, October 2-7.

Hanna, S.R., 1982. "Review of Atmospheric Diffusion Models for Regulatory Applications." World Meteorological Organization Technical Note No. 177, Geneva.

Johnson, J.R. and Dunford, D.W., 1983. "Dose Conversion Factors for Intakes of Selected Radionuclides by Infants and Adults." AECL-7919. Atomic Energy of Canada Limited, Chalk River.

Linsley, G.S., 1978. "Resuspension of the Transuranium Elements - A Review of Existing Data". National Radiological Protection Board, NRPB-R75.

Nordstrom, D.K., 1982. "Aqueous Pyrite Oxidation and the Consequent Formation of Secondary Iron Minerals". In Acid Sulphate Weathering, Soil Science Society of America, 37-55.

Rogers, V.C., Overmeyer, R.F., Putzig, K.M., Jensen, C.M., Nielson, K.K., and Sermon, B.W., 1980. "Characterization of Uranium Tailings Cover Materials for Radon Flux Reduction." Prepared by Ford, Bacon and Davis Utah, Inc., for U.S. Nuclear Regulatory Commission, Report NUREG/CR-1081, FBDU-218-2.

Sehmel, G.A., 1980. "Particle Resuspension: A Review." Environment International, Vol. 4, pp. 107-129.

Silver, M., and Ritcey, G.M., 1982. "A Simulation of the Effects of Bacteria, Organics and Salt Solutions on Uranium Mine/Mill Tailings from Elliot Lake, Ontario". In Waste Treatment and Utilization, Theory and Practice of Waste Management, Vol. 2, M. Moo-Young (Ed.), Pergamon Press, 489-501.

Shumate, K.S., Smith, E.E., Dugan, P.R., Brant, R.A. and Randles, C.E., 1971. "Acid Mine Drainage, Formation and Abatement". E.P.A. Rept. 14010 FPR 04/71.

Stumm, W., and Morgan A.T., 1981. "Aquatic Chemistry". 2nd Ed., John Wiley pub., New York.

Turner, D.B., 1969. "Workbook of Atmospheric Dispersion Estimates." U.S. Dept. of Health, Education and Welfare, Cincinnati, Ohio.

United States Nuclear Regulatory Commission (NRC), 1980. "Final Generic Environmental Impact Statement on Uranium Milling". NUREG-0706.

Zettwoog, P., Fourcade, N., Campbell, F., and Caplan, H.S., 1981. "The Effect of Freezing on the Emanation of Radon from Uranium Ore and Tailings Samples". Health Physics, Vol. 41, pp. 410-411.

APPENDIX D

PARAMETER DISTRIBUTIONS

APPENDIX D
PARAMETER DISTRIBUTIONS

All parameters specified in the component models described in Appendix C were designated either as constant or variable quantities. Parameters assigned variable quantities were described by one of six different parametric distributions depending on which distribution was considered to best describe the data. Some of the input parameters and, of course, the output results, were calculated as functions of the input quantities.

Where appropriate data were available, they were used to estimate an appropriate parametric distribution. A number of input distributions, such as mean wind speed (WIND), fraction of time wind blows to the receptor (FRACT) and net precipitation (NPRCP) were based on extensive empirical data.

In the absence of easily obtainable or appropriate data, the distributions were estimated based on underlying physical principles, trends in similar parameters, and/or professional judgement. A further consideration was that all parametric distributions available in the OPANAL program should be exercised. The six available distributions; normal, lognormal, uniform, log-uniform, triangular and Beta, were outlined in Table 3.1 and Appendix A.

The model parameters, their logical variable names, definitions and associated units are specified in the following table.

KEY TO TABLE D.1

Distribution Type		Distribution Specifications
C	Constant	single numerical value
LN	Log normal	geometric mean geometric standard deviation
N	Normal	arithmetic mean arithmetic standard deviation
T	Triangular	minimum, mode (peak), maximum
U	Uniform	minimum, maximum
LU	Log Uniform	minimum, maximum
B	Beta	minimum, maximum, a, b
calculated		numerical value calculated by model

TABLE D.1

LISTING OF PARAMETERS
WITH THEIR DEFINITIONS, PROBABILITY DISTRIBUTIONS AND UNITS

Symbol	Definition	Used In Model	Parameter Probability Distributions: Distribution Type	Distribution Specifications	Units
ARBASN	area of the tailings basin watershed	$P_{0,2}$, $X_{o,water}$	C	15.0E5	m^2
AREXTL	exposed tailings surface area	$X_{o,air}$	C	50.0E4	m^2
ARLKBT	area of lake bottom	$P_{0,2}$	C	1.6E6	m^2
ARTLSR	total tailings surface area	$X_{o,air}$	C	5.0E5	m^2
BETA	internal variable in model $P_{3,2}$	$P_{3,2}$	Calculated		-
BPRA	water to edible portion of aquatic animal type f transfer coefficient	$P_{2,7}$	LN	15.0E-6;3.44	$Bq.g^{-1}.Bq^{-1}.m^3$
BIOWCL	concentration of aquatic biota in the lake water column	$P_{0,2}$	Calculated		$g.m^{-3}$
DARCY	the Darcy velocity	$P_{3,2}$	Calculated		$m.a^{-1}$
DCFING	dose conversion factor for ingestion	$P_{5,10}$, $P_{6,10}$ $P_{2,10}$, $P_{7,10}$	C	3.2E-7	$Sv.Bq^{-1}$
DCFINH	dose conversion factor for inhalation exposure	$P_{1,10}$	C	2.6E-6	$Sv.Bq^{-1}$
DCFIRN	dose conversion factor for inhalation exposure to radon	$P_{1,10}$	C	7.2E-5	$Sv.a^{-1}.Bq^{-1}.m^3$
DISPER	dispersivity	$P_{3,2}$	N	25;10	m
DISPLT	longitudinal dispersion	$P_{3,2}$	Calculated		$m^2.a^{-1}$
DIST	downwind distance (receptor)	$P_{0,1}$	C	3600	m
DISTG	distance along the groundwater flow path from tailings area	$P_{0,3}$, $P_{3,2}$	C	200	m
EBKDFC	effective bulk diffusion coefficient for radon	$X_{o,air}$	Calculated		$m^2.s^{-1}$
EMANAT	emanation coefficient	$X_{o,air}$	T	0.02;0.2;0.9	-
EMIT	source term for air	$X_{o,air}$	Calculated		$Bq.s^{-1}$

Symbol	Definition	Used In Model	Parameter Probability Distributions: Distribution Type	Distribution Specifications	Units
ESUBV	the fraction of the foliar deposition retained on the edible portions of the plant (0 to 1):				
	a) lichen, summer forage and above ground vegetation	$P_{1,5}$	C	1.0	-
	b) below ground vegetation	$P_{1,5}$	C	0.1	
FACTCO	activity coefficient	$X_{o,water}$	Calculated		-
FDEPVE	the fraction of the total deposition retained on the plant surface (0 to 1):				
	a) lichen	$P_{1,5}$	C	0.95	-
	b) summer forage, above and below ground vegetation	$P_{1,5}$	C	0.20	-
FETOPR	feed-to-product p transfer coefficient for animal type a	$P_{2,6}$,$P_{5,6}$	C	1.4E-9	$Bq.g^{-1}.Bq^{-1}.a$
FNPRCP	fraction of precipitation on the tailings drainage basin that infiltrates and enters the deep groundwater flow below the tailings	$P_{0,2}$,$P_{3,2}$ $X_{o,water}$	C	0.15	-
FRACT	fraction of time the wind blows from source to receptor	$P_{0,1}$	N	0.1;0.02	-
FSUBA	fraction of aquatic animal type f that is contaminated (0 to 1):	$P_{7,10}$	C	1.0	-
FSUBP	fraction of animal produce that is contaminated (0 to 1)	$P_{6,10}$	C	0.5	-
FVEGCN	fraction of vegetation type v consumed by animal type a that is local (0 to 1)	$P_{5,6}$	C	0.5	-

Symbol	Definition	Parameter Probability Distributions: Used In Model	Distribution Type	Distribution Specifications	Units
PVEGGR	fraction of vegetation type v grown at the contaminated source (0 to 1)	$P_{5,10}$	C	0.5	-
FWATHM	fraction of water consumed by a human that is contaminated	$P_{2,10}$	C	1.0	-
GWCONC	concentration of radium in groundwater at discharge to the river	$P_{3,2}$	Calculated		$Bq.m^{-3}$
HSUBD	initial mixing depth	$X_{o,air}, P_{0,1}$	U	1.0;3.0	m
HYDCON	hydraulic conductivity	$P_{3,2}$	N	315;300	$m.a^{-1}$
HYDGRD	hydraulic gradient	$P_{3,2}$	N	0.05;0.01	$m.m^{-1}$
INHALE	annual inhalation rate	$P_{1,10}$	C	8400	$m^3.a^{-1}$
JFLUX	contaminant (radium) flux via the groundwater route to the river	$P_{0,2}, P_{3,2}$	Calculated		$Bq.a^{-1}$
JRAWCO	flux of radium from sediments into the water column	$P_{0,2}$	Calculated		$Bq.m^{-2}.a^{-1}$
KPRIME	ratio of the half saturation constant for oxygen to the oxygen concentration at the tailings surface	$X_{o,water}, X_{o,air}$	LN	0.05;2.0	-
KSUBBI	transfer coefficient for radium from water to aquatic biota	$P_{0,2}$	LN	500E-6;3.2	$m^3.g^{-1}$
KSUBD	sorption or distribution coefficient for radium on suspended solids	$P_{0,2}$	LN	80E-6;3.5	$m^3.g^{-1}$
KSUBDG	distribution coefficient for radium on geologic material along groundwater path	$P_{3,2}$	N	25E-6;10E-6	$m^3.g^{-1}$
KSUBL	mass transfer coefficient of radium into the lake water column from sediments	$P_{0,2}$	LN	0.5;2.24	$m.a^{-1}$
KSUBSO	solubility equilibrium constant for gypsum	$X_{o,water}, P_{0,2}$	C	25.0	$moles^2.m^{-6}$

Symbol	Definition	Parameter Probability Distributions: Used In Model	Distribution Type	Distribution Specifications	Units
LAMRA	radioactive decay constant for radium	$P_{3,2}$	C	430E-6	a^{-1}
LAMRN	radioactive decay constant for radon	$X_{o,air}$	C	2.1E-6	s^{-1}
LAMWTH	a decay constant accounting for weathering losses	$P_{1,5}$	C	5.73E-7	s^{-1}
LSUBC	radon diffusion length in the cover	$X_{o,air}$	U	0.5;1.5	m
LSUBT	radon diffusion length in tailings	$X_{o,air}$	Calculated		m
MPCH2O	percent water (by weight) in tailings	$X_{o,air}$	T	5.0;15.0;25.0	-
NPRCP	net precipitation (i.e. after evapotranspiration losses)	$X_{o,water}$, $P_{0,2}$ $P_{3,2}$	N	0.4;0.1	$m.a^{-1}$
NSUBW	number of winter months per year	$X_{o,air}$	C	5.0	-
OCCPAC	occupancy factor, or fraction of year spent at the location (0 to 1)	$P_{1,10}$	T	0.3;0.6;0.9	-
O2CONC	oxygen concentration at the the tailings surface	$X_{o,water}$	C	9.04	$moles.m^{-3}$
O2DIFF	oxygen diffusivity through the tailings	$X_{o,water}$	LN	70.0;1.35	$m^2.a^{-1}$
PHI	concentration of sulphate in porewater	$X_{o,water}$	Calculated		$moles.m^{-3}$
POROSG	porosity of the geologic material	$P_{3,2}$	N	0.3;0.025	$m^3.m^{-3}$
PROBIO	net biomass productivity	$P_{0,2}$	LN	30.0;1.15	$g.m^{-2}.a^{-1}$
QSUBI	flowrate of the freshwater in the river at the point of discharge to the lake excluding the tailings sub-basin discharge	$P_{0,2}$	Calculated		$m^3.a^{-1}$
QSUBO	total water outflow from the lake	$P_{0,2}$	Calculated		$m^3.a^{-1}$
QSUBR	total flow in river system at end of reach with discharge to the lake including the tailings sub-basin discharge	$P_{0,2}$	Calculated		$m^3.a^{-1}$

Symbol	Definition	Used In Model	Parameter Probability Distributions: Distribution Type	Distribution Specifications	Units
QVEG	annual intake of vegetation type v (dry wt. basis) by animal type a:	$P_{5,6}$			$g.a^{-1}$
	a) lichen by caribou		C	1490E3	$g.a^{-1}$
	b) summer forage by caribou		C	1065E3	$g.a^{-1}$
	c) above and below ground vegetation		-	0.0	-
QVOLSR	volumetric rate of groundwater flow from the tailings area	$P_{3,2}$	Calculated		$m^3.a^{-1}$
RAAIR	airborne concentration of radium	$P_{0,1}$	Calculated		$Bq.m^{-3}$
RACNTA	radium concentration in the tailings	$X_{o,air}$	U	5.0;15.0	$Bq.g^{-1}$
RAGPTA	radium content of gypsum in tailings	$X_{o,water}$	B	2500;10,000;a=1.5;b=.5	$Bq.mole^{-1}$
RALTOT	concentration of total radium in the lake water column	$P_{0,2}$	Calculated		$Bq.m^{-3}$
RAPOWA	radium concentration in the tailings porewater	$X_{o,water}, P_{0,2}, P_{3,2}$	Calculated		$Bq.m^{-3}$
RARTOT	concentration of total radium in the river at its point of discharge into the lake	$P_{0,2}$	Calculated		$Bq.m^{-3}$
RATI	concentration of radium in the river system upstream of the tailings discharge	$P_{0,2}$	C	37.0	$Bq.m^{-3}$
RESDEP	effective resuspension depth	$X_{o,air}$	U	1.0E-3;1.0E-2	m
RESFAC	resuspension factor for dust from tailings	$X_{o,air}$	LU	1.0E-9;1.0E-5	m^{-1}
RETFAC	retardation factor	$P_{3,2}$	Calculated		-
RHOBAS	bulk density of the geologic material	$P_{3,2}$	N	1.92E6;0.15E6	$g.m^{-3}$
RHOTAL	tailings density	$X_{o,air}, X_{o,water}$	N	1.4E6;0.1E6	$g.m^{-3}$
RNSOUR	source emission rate for radon	$X_{o,air}$	Calculated		$Bq.s^{-1}$
SIGBIO	sedimentation velocity of biomass	$P_{0,2}$	LN	12.0;1.15	$m.a^{-1}$
SIGSS	sedimentation velocity of suspended solids	$P_{0,2}$	LN	24.5;1.1	$m.a^{-1}$
SIGZED	vertical dispersion coefficient	$P_{0,1}$	LN	60.0;1.73	m

Symbol	Definition	Used In Model	Parameter Probability Distributions: Distribution Type	Distribution Specifications	Units
SSL	concentration of suspended solids in the water column	$P_{0,2}$	Calculated		$g.m^{-3}$
SSR	suspended soils concentration in the river at its point of discharge to the lake	$P_{0,2}$	LN	4.0;2.0	$g.m^{-3}$
SULPRD	average yearly sulphate production rate	$X_{o,water}$	LN	1.4E-6;1.73	$moles.(g.a)^{-1}$
TEXPDP	duration of plant exposure for vegetation type v to atmospheric deposition:				
	a) lichens	$P_{1,5}$	C	100.0E6	s
	b) summer forage	$P_{1,5}$	C	2.592E6	s
	c) above and below ground vegetation	$P_{1,5}$	C	5.184E6	s
TIME	time from placement of tailings	$P_{3,2}$	U	50;2500	a
THICK	clean cover thickness : Option #1	$X_{o,air}$	C	0.0	m
	Option #2	$X_{o,air}$	T	0.5;1.0;2.0	m
			T	0.5;1.0;2.0	
USUBA	annual human ingestion rate of aquatic animal type f	$P_{7,10}$	C	8000	$g.a^{-1}$
USUBP	the annual human ingestion rate of animal produce type p	$P_{6,10}$	C	79.0E3	$g.a^{-1}$
USUBV	the annual human ingestion rate of vegetation type v:				
	a) lichen and summer forage	$P_{5,10}$	C	0.0	$g.a^{-1}$
	b) above ground vegetation	$P_{5,10}$	C	157.0E3	$g.a^{-1}$
	c) below ground vegetation	$P_{5,10}$	C	70.0E3	$g.a^{-1}$
USUBW	annual human intake of water	$P_{2,10}$	C	0.6	$m^3.a^{-1}$
VELDEP	deposition velocity	$P_{1,5}$	C	0.02	$m.s^{-1}$
WIND	wind speed	$X_{o,air}$, $P_{0,1}$	N	5.3;0.5	$m.s^{-1}$

Symbol	Definition	Used In Model	Parameter Probability Distributions: Distribution Type	Distribution Specifications	Units
WTSPLX	winter to summer radon flux ratio	$X_{o,air}$	LN	0.02;2.15	-
XOGWRA	source term release of radium to the groundwater flow component	$X_{o,water}$	Calculated		$Bq.s^{-1}$
XOSWRA	source term release of radium to the surface water flow component	$X_{o,water}$	Calculated		$Bq.s^{-1}$
YDENST	the yield density of vegetation type v:				
	a) lichen	$P_{1,5}$	C	500.0	$g.m^{-2}$
	b) summer forage	$P_{1,5}$	C	750.0	$g.m^{-2}$
	c) above and below ground vegetation	$P_{1,5}$	C	2000.0	$g.m^{-2}$
ZETA	concentration ratio for dissolved radium between interstitial (sediment) water and the lake water.	$P_{0,2}$	U	0.5;1.5	-

APPENDIX E

DISCUSSION PAPERS

E.1 Probabilistic Endpoint for Assessment

E.2 Sensitivity Analysis for Environmental Simulation Models

E.3 Parameter Uncertainty Analysis Application

E.4 Alternative Sampling Procedures

E.5 Software Quality Assurance: Current Concepts

APPENDIX E.1

The Probabilistic Endpoint for an Assessment

Prepared by

R.V. Osborne

For

SENES Consultants Limited

June 1984

1.0 Basic Principles

The starting point for an assessment of the radiological factors that will constrain how any radioactive waste is managed or disposed is the ICRP "system of dose limitation". Paragraph (12) of ICRP 26 [1] reads:

> "... the Commission recommends a system of dose limitation, the main features of which are as follows:
>
> (a) no practice shall be adopted unless its introduction produces a positive net benefit;
>
> (b) all exposures shall be kept as low as reasonably achievable, economic and social factors being taken into account: and
>
> (c) the dose equivalent to individuals shall not exceed the limits recommended for the appropriate circumstances by the Commission".

The last feature (or principle) is meant to ensure that, in applying the second principle (the ALARA principle) no individual incurs an unacceptable exposure. For this purpose, the concept of a critical group in an exposed population has been introduced. The dose limits refer to the average individual in the critical group.

On the basis of a comparison with the levels of risk regularly accepted in everyday life, the ICRP concluded that a level of fatal risk from justified radiation exposures of less than about 10^{-5} per year would be likely to be acceptable to any individual member of the public. Risk of 10^{-5} per year

corresponds to about 0.5-1 mSv per year of whole body exposure.

For practices or activities which are unlikely to give rise to a persistent level of exposure, the ICRP recommends an effective dose equivalent limit of 5 mSv per year. According to the ICRP, the application of such a limit is likely to result in average dose equivalents to individuals of the public not higher than 0.5 mSv per year provided that the practices exposing the public are few and cause little exposure outside the critical groups. It is considered by the ICRP to guarantee an appropriate degree of protection, even if the few individuals exposed near to the dose limit could run a higher risk, in the range of 10^{-5} to 10^{-4} per year. It is however noted by the ICRP (paragraph 122, loc. cit) that where exposures to a few individuals could persist at high rates over prolonged periods, it would be prudent to take measures to restrict their lifetime dose to less than 1 mSv per year.

The imposition by national regulation of an annual dose limit for individual members of the general public follows from these ICRP recommendations. Application to individuals exposed as a result of operating facilities is generally straightforward; the exception is the application to individuals exposed to major accidental releases. Although the distinction is really one of degree, in practice a sharp distinction is made between determining that exposures of individual members of the public are below a given limit when the fact of exposure is considered very likely and when the exposure is considered to be unlikely.

In a similar way, in determining whether long-lived radioactive waste is disposed of acceptably, not only do the magnitudes of possible exposures need to be considered but also the probabilities that these exposures will arise. This is because scenarios may often be defined that result in exposures above any single limit but which are very improbable.

The general idea, therefore, in protecting individuals, is that a limit should be placed on a quantity that combines the probability of an exposure occurring with the dose from such an exposure. This is loosely referred to as a "risk"

limit.

2.0 The Risk Limit

Since the word risk is used in several ways, some clarification of what is intended here is needed. If the consequence of an event is represented by C, and the probability of that event occurring is p, then the statistical expection, E, is given by:

$$E = p \times C$$

Each of E, p and C has been referred to as 'risk" at various times - hence the considerable confusion. Here we are concerned with the product of p and C. If we take the probability of a fatal effect from radiation (r) as 10^{-2} Sv^{-1} and probability of a dose D Sv being received as p, then the value of E is given by:

$$E = rpD$$

If we take r as fixed by definition, then a "risk" limit could be in terms of E, the expectation of death, or in terms of (E/r), the expectation of dose. The current annual limit on dose is a special case of the latter with p = 1. As noted above, the appropriate annual limit on dose is 1 mSv when exposures may be expected to continue for much of any individual's lifetime. Therefore the annual limit on expectation of dose could also be 1 mSv. There is a value judgement in such a conclusion; for present purposes we can take it as tenable. That being so, the limit for E/r would be met if:

$$pD < 10^{-3}\ Sv.a^{-1}$$

where D is the annual dose and p is the probability of it being received. The equivalent limit on E is given by:

$$rpD < 10^{-5}\ a^{-1}$$

The above limit is appropriate when the doses being considered are sufficiently low that there is no possiblity of acute health effects, the effects usually termed "non-stochastic". Some modification of the "risk" limit above is clearly needed when the doses reach the "non-stochastic" region. The limiting probability of any value of dose in this high range should be independent of the value of dose. An extreme example is that a dose of 10 Sv should be just as unlikely as a dose of 100 Sv.

One can regard such acute effects as 'certain' effects and limit the probability of such doses to the limiting value of E, i.e. 10^{-5} per year. The appropriate value for D at which this transition in limit should occur is subject to debate: the NEA report [2] recommends an annual dose of 1 Sv. With this value, the "risk" to average members of a critical group is limited by specifying:

$$pD < 10^{-3}\ \text{Sv.a}^{-1}$$
$$\text{for } D < 1\ \text{Sv.a}^{-1},$$
$$\text{and } p < 10^{-5}$$
$$\text{for } D \geq 1\ \text{Sv.a}^{-1}$$

Implicit in this statement of the limit is that if D is less than 1 mSv.a^{-1} then the probability of the D being received does not matter for comparison with a limit.

The direction of an assessment must therefore be to ascertain for members of any critical group whether:

(a) annual doses are less than 1 mSv.a^{-1}, or

(b) the expectation of annual dose is less than 1 mSv.a^{-1} if the annual dose is less than 1 Sv, and

(c) the probability of an annual dose greater than or equal to 1 Sv is less than 10^{-5}.

The above discussion has considered a single 'event' with which a particular

dose is associated and which has a particular likelihood of occurring. It is not self-evident how these ideas are to be employed in an actual assessment. Further argument is needed as follows:

3.0 Probabilistic Assessments

Probabilistic ideas can be introduced into an assessment in several ways.

i) We can estimate the annual dose (D_c) to a critical group associated with a particular pathway in the traditional way, using "reasonably conservative" models and parameters. The probability of transport of a radionuclide along that particular pathway might be judged to be less than unity - say p_c. The expected annual dose is then $P_c D_c$. An example is the dose that might occur as a result of a particular kind of farming close to a waste site. The probability of that kind of farming occurring might be judged to be some value less than one.

ii) The annual dose by a particular pathway may be estimated using distributions of values for the parameters. [These distributions may reflect actual observations or they may reflect a lack of knowledge of what the values may be]. The probability of that particular pathway may be judged to be some value less than unity. The problem is then to combine these two probabilities.

There are two possibilities:

(a) Because we have already defined the critical group we are interested in the expectation of the annual dose to that group. If the estimated annual dose distribution were p(D) for this group by this particular pathway, then

$$\int_0^\infty p(D)\, dD = 1$$

and

the expectation of the annual dose if the pathway occurs is

$$\int_0^\infty p(D)\ D\ dD = \overline{D} \text{ say}$$

For example, we may have assumed some distribution of the amount of fish eaten by an average individual in a critical group that is associated with the fish ingestion pathway. We would carry out our analysis with the mean value for the amount of fish eaten. If the probability of transport along this pathway is P_c as before, then the expected annual dose is $P_c\ \overline{D}$.

(b) Rather than taking the mean dose, D, by that particular pathway, we might choose to select a value in the upper part of the estimated distribution, say the value that is exceeded only 10% of the time ($D_{0.9}$). The justification would be that this represents a more reasonable selection of 'conservative' parameters for the critical group. The expected annual dose for comparison with the probabilistic dose limit would be $P_c\ D_{0.9}$. It should be noted that this method is really a redefinition of the critical group. The annual dose $D_{0.9}$ is the expectation value for a sub-set of the original group that is associated with only part of the distributions of various parameters. If we use the same example as above, then we have, in effect, redefined our critical group; it now has the attribute of individual fish consumption much greater than that of the original critical group. It is obviously a matter of judgement as to whether this new de facto definition of critical group is to be preferred. It might be argued that since our "best" subjective judgment went into the selection of the critical group in the first place we are being inconsistent in changing the selection in this iterative way.

iii) Several pathways (or scenarios) may be considered in an assessment of the doses received by one or more critical groups. [Although strictly there is only one critical group for a given source of radionuclides, in practice many groups, associated with different pathways, are likely to

be considered in an assessment. Each is potentially the critical group]. We may consider two possibilities:

(a) The critical groups are such that no one individual could be a member of more than one group. The expectation of the annual dose may then be estimated as in 1 or 2 above for each critical group and pathway. An example is the pair of critical groups, one that is assumed to obtain all its food from local agriculture, the other that is the hunting/fishing group on a lake by a waste site.

(b) A critic[illegible]roup may have members that may be exposed to irradiation as a result of more than one pathway. If the annual dose D_i is expected as a result of pathway i, and if the probabilities are independent, then the expectation of the annual dose is $\Sigma P_i D_i$ where P_i is the likelihood of pathway i occurring. [From (1) and (2) above, D_i would be D_c, D or $D_{0.9}$ in the previous nomenclature]. An example is when the critical group resides close to a waste site and has a critical pathway by fish ingestion if radium, say, is leached into the aquatic pathway or by inhalation of radon and daughters if the radium remains in the waste. This assessment becomes much more complicated when the various pathways are not exclusive and not independent. If the probabilities of the doses by the various pathways actually being received are small compared with unity, then the non-exclusiveness does not matter and the simple summation above can be used.

4.0 Time Variation

In the above, variations of probabilities and annual doses with time have been ignored. An implicit assumption has been that all the annual dose distributions calculated do not vary over the lifetime of individuals in the critical group(s). For example, in the fish pathway we note a distribution in estimated annual dose arising because of a distribution in annual intake of fish. However, the annual dose as a result of such a pathway is likely to

vary over time, perhaps as the release from a waste site varies over hundreds of years. Clearly, it is not valid to take the expectation of such a distribution as the appropriate annual dose to use in the type of estimate noted in (1) and (2) above because a single individual is not involved with the total distribution over time. It is appropriate here to determine the maximum expectation of annual dose for a particular pathway averaged over a 50-year period (or 70-year period; essentially a lifetime). The probability of that pathway occurring is used with such an expectation as in (2) above.

When several pathways are involved each resulting in different distributions of annual dose in time then a simple combination of expectations as in 3(b) is inappropriate. This is because a single individual may not be at risk from the maximum expectations of annual doses from each of the pathways. Therefore, concurrent annual doses (or distributions of annual doses) averaged over 50-year periods, by the various pathways need to be estimated. The variation in time of the expectation of the annual dose (averaged over 50-year periods) may then be obtained. For example, suppose that we can assign likelihood P_1 to circumstances that will result in a 50-year average annual dose of D_1 to an individual in a critical group over the period 200-500 years after closure of a waste site and P_2 to other circumstances that result in a similar dose, D_2, over the period 600-1000 years. The values of P_1D_1 and P_2D_2 should be compared independently with the 'risk' limit since different individuals must be involved. If the latter dose could be received by the same critical group in the period 400-800 years, then the expectation of annual dose through the period 400-500 years would be $P_1D_1 + P_2D_2$ and would be the appropriate quantity to compare with the 'risk' limit.

5.0 Estimation of Probabilities

The assignment of probability distribution to parameters used in assessment models can, in principle, be based on observations although even with readily observed pathways - for example, foodchains - considerable judgement is needed if predictions are being attempted for times in the distant future.

In fact, many of the probabilities that we are dealing with are subjective rather than objective. For the purposes of an assessment we can treat them according to probability theory. However, their subjectivity should temper any attempt to assign precision to the numerical outcome of an assessment.

The assignment of probabilities to the occurrence of particular pathways may, sometimes also, be extrapolated from observations. For example, different kinds of farming (resulting in different foodchains) may be considered each to occur in a given area (close to a waste site) over less than an individual's lifetime. The probabilities for pathways associated with the different kinds may then be derived from current land-use statistics for the area. For some pathways the estimates of probabilities may be very subjective. In these instances a more appropriate approach may be to assign unity probabilities to the pathway and then determine how unlikely that pathway has to be for the expectation of the annual doses to be acceptable. The judgement needed is then whether such a probability is reasonable.

In the above example, the qualification that each of the different kinds of farming had to occur within a lifetime (or, say, 50 years) is important. If it were considered that any type might persist for 50 years or more then the critical pathway would be associated with a single type of farming which would have unity probability.

For many pathways, assignment of unity likelihood is appropriate anyway since there will be little doubt that they will occur. An example is the radon and daughters inhalation pathway. The assumption usually made is that a critical group is present to receive a radiation dose by a particular pathway. That is, no credit is taken for the probability that a critical group may not be present at the appropriate location - or does not have a particular characteristic. For example, if fish in a lake may be contaminated by radionuclides from a nearby waste site it is usually assumed that individuals (defined in this way as the critical group) are present and do eat the fish. However, judging where is the appropriate location for individuals exposed to airborne radon from tailings is clearly very subjective. Do they live at the

edge of the tailings or 1 km away? Although some probability distribution for the separation of a critical group from the tailings may be defined it is not obvious on what it could be based.

Some doses may occur as a result of particular events to which is ascribed an annual probability of occurrence. Examples are dam failure, which allows tailings to be distributed, and intrusion into a waste site. Different groups have taken different approaches to using such annual probabilities with widely different results (see [2]). If the annual dose is D after the event occurs then one approach is to take the expectation of the annual dose as $P_a D$ where P_1 is the probability of the event per year. This approach clearly has dimensional difficulties and yields low values of expectation. A second approach is to argue that over a long enough time the likelihood of the event approaches unity so, since the critical group can be placed whenever the event occurs the expectation of the annual dose is D. This estimation can readily include the effect of a variation with time of the annual probability of the event occurring and the reduction with time of D because of radioactive decay. Nevertheless the approach given high estimates of expectation of annual dose since no one individual is at risk over a very long time. A better approach is to take the probability of the event occurring in a 50-year period (or lifetime) as the appropriate value of probability of annual dose D being received by any one individual. The expectation of annual dose is then $50\ P_a D$ if P_a and D are constant in time.

References

[1] International Commission on Radiological Protection, Recommendations of the International Commission on Radiological Protection Annals ICRP 26, 3 (1977).

[2] Nuclear Energy Agency, Long Term Radiation Protection Objective in Radioactive Waste Disposal, NEA(OECD) Report, in press (OECD, Paris).

Appendix E.2

Sensitivity Analysis for Environmental Simulation Models: Problems, Definitions and Current Opinion

Prepared by
Robert H. Gardner and F. Owen Hoffman

for
SENES Consultants Limited
June, 1984

1.0 Introduction

The usefulness of simulation methods for the synthesis of information and the prediction of future conditions has led to their proliferation for solving environmental problems. The development and use of complex models requires a simple and rigorous method of evaluating the effect that model components and parameters have on model predictions. These methods are often collectively referred to as "sensitivity methods", "sensitivity analysis", or "model sensitivities."

Although for most modelling studies these terms have been (or can be) rigorously defined, they are often freely used without precise definition. For instance, questions concerning the sensitivity of model predictions to model assumptions seems to be a legitimate use of the word "sensitivity", yet methods for quantifying the effect of different assumptions on model predictions are very difficult to define and quantify in a formally consistent manner. On the other hand, a variety of techniques have been developed to determine the sensitivity of model predictions to values selected for input. However, precise definitions of the technique and the question it is designed to address are required before the term "sensitivity analysis" can be clearly interpreted.

2.0 Tomovic Sensitivity and the Slope of the Line

Classical sensitivity methods were first proposed by Tomovic (1963) as the rate of change of the state of the system (X) as a result of a change in a particular parameter (P). This problem is easily solved for time-invariant parameters and static systems, whether linear or nonlinear, by simple calculus (the partial derivative of X with respect to P). (NOTE: The model need not be linear, only the response of the model in the small region around the nominal value of the parameters.) The primary uses of Tomovic sensitivity are to explore responses to small changes in inputs, initial conditions, and parameters; reveal functional relationships within the model; and test the local stability of model behaviour near some equilibrium point. For these reasons, Tomovic sensitivity is usually preferred in mathematical and engineering applications.

The problems with Tomovic sensitivity, at least in environmental simulations are: (1) the equations, although conceptually simple, are extremely difficult to develop. As a result, numerical solutions (i.e., computer perturbation experiments) are usually required; (2) Tomovic sensitivity measures linear responses of the model to small changes in individual parameters. Therefore, Tomovic sensitivity is not useful when parameters are simultaneously varied over large ranges, and often gives incorrect values in these situations (Gardner et al, 1981). These difficulties have led to the development and use of a variety of techniques for model evaluation (Rose and Swartzman, 1981).

3.0 Statistical Measures of Sensitivity

If classical sensitivity methods are concerned with the effect of P (a model parameter) on the local dynamics of X (a model response or prediction) then "statistical sensitivity" is concerned with the effects of imperfectly known P's (parameters with associated variances) on the confidence level of model predictions (variations in the X's). "Statistical sensitivity" studies have been referred to as error analysis or uncertainty analysis to distinguish them from classical sensitivity analysis (O'Neill and Gardner, 1979). Some

questions addressed by these methods are: (1) what is the confidence level for model predictions? (2) which parameter is most important in determining the uncertainty of model outputs? and (3) is the variance of model predictions reduced by more precise parameter estimation?

The advent of large computers and the development of Monte Carlo simulation methods made possible the statistical characterization of many environmental models. Monte Carlo methods involve the iterative selection of random values for the model parameters from specified frequency distributions, performing a model simulation to obtain the predictions that are a function of that random set of parameters, and analyzing the combined set of parameters and predictions (McGrath et al, 1975, Rubinstein 1981). Systematic sampling methods for Monte Carlo simulations are more efficient than simple random sampling because (1) the estimates of output parameters (e.g., mean, median, mode) are more precise (see McKay et al, 1979), (2) low rates of spurious relationships between randomly generated values are ensured (Iman and Conover, 1982), and (3) computer codes exist for generating values from a variety of prespecified distributions (Gardner et al, 1984). When variances are small (<1% of the nominal value) then statistical sensitivity methods and Tomovic sensitivities are identical (Gardner et al, 1982; Gardner in press).

The data set produced from Monte Carlo simulations (e.g., sets of P's with corresponding sets of X's) allows the use of many different measures of dependencies of the X's on the P's. Many unique methods have been suggested (Rose and Swartzman, 1981), but for general use the best measures are those based on the theory and rigor of correlation and regression procedures. Simple correlation coefficients and more complex multiple regression methods have been extensively used to indicate the degree of association between parameter uncertainty and prediction uncertainties (for more details see Downing et al).

4.0 Conclusion: Our Position

Many techniques are available for numerical analysis of model dynamics (Table

1). Our preferences are for those that are simple to use, general in their application, and clear in their interpretation. However, the limiting factor in the analysis of most assessment models is the availability of adequate information for estimating parameter values and describing their probabilistic behaviour. When little information is available (other than the nominal value of the parameters) then Tomovic sensitivity is useful for model screening. If a parameter has a high Tomovic sensitivity and large uncertainty, then this parameter will contribute significantly to model uncertainties. If the parameter has a low sensitivity value and low uncertainty, then this parameter will contribute little to prediction uncertainties. The problem arises when parameters have a low sensitivity and high uncertainty, or vice versa. In these cases, Tomovic sensitivity is of limited use (Hoffman and Gardner, 1984).

It is difficult to obtain adequate estimates of nominal values and associated frequency distributions for all parameters of a model. Thus, most assessments are performed with less than adequate information. In these cases, simple measures of dependency between predictions and parameters appear to be most useful. Spearman correlation coefficients (the simple correlation of the rank transformed data) are useful because they are easily calculated from the data provided by Monte Carlo simulations, they are not sensitive to departures from linearity, and they provide a reliable ranking of parameter importances. As more information becomes available, multiple regression methods can be used for more precise determination of the effect of each parameter on prediction uncertainty (Downing et al).

Table 1 is an attempt to synthesize some of the diverse aspects of this problem and the techniques often used to solve them. The intention of this table is to be informative and provocative, but certainly not dogmatic! The rows of the table loosely correspond to increasing levels of information for describing model parameters. The three sensitivity methods either analytically (Tomovic) or numerically estimate sensitivities by perturbing parameters and simulating the model. The statistical methods (Spearman Correlation and Regression Methods) assume the availability of a data set

generated by some form of Monte Carlo simulations of the model. (The efficiency of various Monte Carlo methods is beyond the scope of this paper). Factorial methods can be used to generate a data set similar to the Monte Carlo methods (hence the (?) symbol in the table), but this is rarely done because more iterations are needed for a complete factorial than a simple Monte Carlo when the number of model parameters is greater than 12.

Sensitivity methods are unable to use information which describes the statistical properties of model parameters - and they are, therefore, unsuitable for uncertainty analysis. Because the sensitivity coefficient can be estimated by regression methods (Gardner in press), these methods are the most general of all situations listed in the Table 1.

Table 1

Some Model Evaluation Methods and Their Relative Usefulness[1]

	Sensitivity Methods			Statistical Methods	
Evaluation Criterion	Tomovic	One at at Time	Factorial Sampling	Spearman Correlation	Regression Methods
Model Evaluation:					
Stability	+++	+	+	+	++
Functional Behaviour	+++	+++	+++	+++	+++
Evaluating Parameters:					
Screening	++	0	0	0	0
Nominal Value	++	++	++	+++	+++
Variance	0	0	?	+++	+++
Distribution	0	0	?	++	+++
Absolute ranks	0	0	?	+	+++
Relative ranks	+	+	?	+++	+++
Large Numbers	+	+	?	++	+++

Tomovic Sensitivity: defined in text; One at a Time: single perturbation experiments; Factorial Design: systematic perturbation experiments so all combinations of parameter values are sampled (McKay et al 1979); Spearman Correlation and Multiple Regression methods are defined in text.

Evaluation criterion: Stability - behaviour of model near an equilibrium point; Functional Relationships - relationships between model mathematics and prediction; Screening - initial estimates of parameter importances; Nominal Value - response of the model to slight changes in nominal values of parameters; Variance - response of model to levels of parameter variance; Distribution - evaluation of importance of shape of distribution; Absolute importance - estimation of the magnitude of effect of each parameter on the response; Relative importance - percent effect of parameters on the response; Large numbers - ability of technique to evaluate models with many parameters.

Key to symbols: (0) not useful: (+) sometimes useful if certain restrictions are met; (++) generally useful with few limitations; (+++) always useful; (?) usefulness depends on factors external to technique (see text).

References

Gardner, R.H., O'Neill, R.V., Mankin, J.B., and Carney, J.H., 1981. "A Comparison of Sensitivity Analysis and Error Analysis Based on a Stream Ecosystem Model". Ecological Modelling, Vol. 12, pp. 177-194.

Gardner, R.H., Rojder, B., and Bergstrom, U., 1983. "PRISM: A Systematic Method for Determining the Effect of Parameter Uncertainties on Model Predictions." Studsvick Energiteknik AB report, Studsvik, Sweden.

Gardner, R.H. in press. "Error Analysis and Sensitivity Analysis in Ecology." Encyclopedia of Systems and Control, Madan Singh, ed., Pergamon Press, London.

Gardner, R.H., 1984. "A Unified Approach to Sensitivity and Uncertainty Analysis." Proceedings of ISTED International Symposium, ASM '84, June 4-6, San Francisco. [Copy Appended]

Downing, D.J., Gardner, R.H. and Bartell, S.M., Manuscript. "Univariate and Multivariate Methods for Analysis of Sets of Predictions from Monte Carlo Simulations". Submitted to Communications in Statistics.

Hoffman, F.O. and Gardner, R.H., 1984. Evaluation of Uncertainties in Environmental Radiological Assessment Models." In: Radiological Assessment, J.E. Till (ed.), U.S. Nuclear Regulatory Commission.

Iman, R.L. and Conover, W.J., 1982. "A Distribution-free Approach to Inducing Rank Correlation Among Input Variables for Simulation Studies." Communications In Statistics B11(3).

Iman, R.L., Davenport J.M., and Zeigler, D.K., 1980. "Latin Hypercube Sampling (Program User's Guide). SAND79-1473, Sandia National Laboratory, Alburquerque, NM.

McGrath, E.J., Basin, S.L. Burton, R.W., Irving, D.C., Jaquette, S.C. and Ketler, W.R., 1975. "Techniques for Efficient Monte Carlo Simulation. Vol. 1: Selection Probability Distributions." ORNL/RSIC-38. Oak Ridge National Laboratory, Oak Ridge, Tennessee. 163 pp.

McKay, MD., Beckman, R.J. and Conover, W.J., 1979. "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code." Technometrics 21: pp 239-245.

O'Neill, R.V. and Gardner, R.H., 1979. "Sources of Uncertainty in Ecological Models." In: B.P. Zeigler, G.J. Kliv, T.I. Oren (eds), Methodology in Systems Modelling and Simulation. pp. 447-463. North-Holland Publishing Co.

Rose, K.A., and Swartzman, G.L., 1981. "A Review of Parameter Sensitivity Methods Applicable to Ecosystem Models." NUREG/CR-2016-2016. U.S. Nuclear Regulatory Commission, Washington, D.C.

Rubinstein, R.Y., 1981. "Simulation and Monte Carlo Method." John Wiley and Sons, New York.

Tomovic, R., 1963. "Sensitivity Analysis of Dynamic Systems." McGraw-Hill Book Co., Inc., New York, 142 pp.

Appendix E.3

Parameter Uncertainty Analysis Application
-- A Position Paper --

Prepared by
F. Owen Hoffman and Robert H. Gardner

for
SENES Consultants Limited
July 198[illegible]

1.0 Introduction

"Parameter uncertainty analysis" is defined as the determination of the uncertainties in the predictions of a mathematical model that are a function of uncertainties in the estimation of the model's parameters (Hoffman and Gardner 1983). The analysis is most often used for the evaluation of "deterministic" mathematical models (i.e. models whose parameters are fixed at a single value for a given prediction). Parameter uncertainty analysis is conducted as follows: 1) The uncertainty associated with the estimation of each model parameter is represented by a distribution of possible values. 2) Statistical error propagation formulae or numerical simulation techniques use these parameter distribution with the model to produce a distribution of predicted values (Fig. 1). 3) The importance of a parameter is determined by the degree to which its prescribed distribution influences the distribution of model predictions.

We caution, however, that sources of uncertainty, such as bias in the structure of the model and bias in the estimation of parameter distributions, are not explicitly included within this type of analysis. The analysis of these additional sources of error is preferably performed by testing model predictions against independent sets of observations (model validation) or by comparing the predictions of different models. In the absence of model

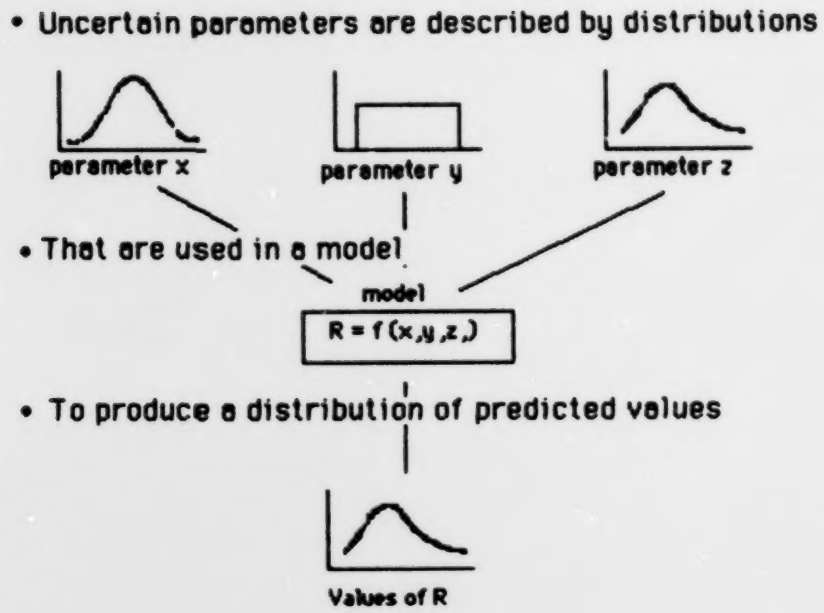


Figure 1. A diagramatic sketch of the initial steps in a parameter uncertainty analysis

validation and model comparison studies, decisions based on the results of a parameter uncertainty analysis must at least qualitatively consider the potential significance of these sources of systematic error.

In order to compensate for the effects of systematic errors and to encompass the range of values that are possible for an uncertain parameter, conservative assumptions are usually employed. Varying degrees of conservative bias are often built intentionally into the structure of the model, and the distributions prescribed to represent parameter uncertainty will usually have a much wider spread than distributions associated with natural variability. Thus, the true probability of an actual value being less than or equal to a predicted value is likely to be much larger than the cumulative percentile determined from the distribution of model output.

Despite the sources of conservative bias which may be built into a model, no model can give an absolute guarantee that values exceeding its predictions will not occur. At best, models are only inexact representations of the real world. There is always a possibility that important processes have been overlooked and that a system at future points in time will behave differently than it has in the past.

2.0 Using the Distributions Produced by Parameter Uncertainty Analysis for Decision-Making

How should decisions be made once a distribution of predicted values is produced? Should one be concerned with the maximum predicted value, some confidence interval about the expected value, or some other indicator of predictive uncertainty?

For the reasons mentioned in the previous section, we find it most productive to focus our attention on a "probable range" of predicted values. This is a range that is reasonably likely to encompass a true value. Meaningful estimates of maximum or extreme values in the tails of predicted distributions are currently beyond the capabilities of extant models and data sets.

We estimate a "probable range" by calculating a 95% confidence interval "of sorts". The qualifying term "of sorts" is used to indicate that the interval is an approximate estimate of a range beyond which values will have a low probability of occurrence. Thus, the 95% confidence interval "of sorts" has more of a qualitative than quantitative significance. The chance of actual values being equal to or less than the maximum value of this interval should be much greater than 97.5%, due to the influence of the above mentioned sources of conservative bias.

The interval can be estimated conveniently using classical statistical methods for calculating a 95% confidence interval for a lognormal distribution: i.e. The maximum value of the interval is obtained by multiplying the geometric mean by the square of the geometric standard deviation, and the minimum value is obtained by dividng the geometric mean by the square of the geometric standard deviation. This calculation is preferred to estimating a 95% interval directly from the predicted distribution because the number of computer iterations required to produce precise estimates of the geometric mean and geometric standard deviation is much less than the number required for precise estimation of the 2.5th and 97.5th cumulative percentiles of the predicted distribution. Our choice of the 95% confidence interval is arbitrary. This interval, however, is one that is commonly used when making decisions about environmental data.

The lognormal distribution is selected because multiplicative chain models composed of distributed parameters will tend to produce a distribution of predicted values that approximates the lognormal even if the distributions of the individual parameters themselves are not lognormal (Aitchinson and Brown 1967). In addition, distributions produced by radiological assessment models and distributions of doses observed within human and animal populations are most prevalently described by (or at the very least encompassed by) the lognormal distribution.

Decisions can be based on a calculated probable range of values in the

following manner:

i) If a limiting value representing a level of concern to a decision-maker is greater than the maximum value of the probable range, no further evaluations should be necessary. Reasonable assurance is provided that a level of concern will not be exceeded, despite the extent of the uncertainty in the model's prediction.

ii) If the limiting value approaches, but is not greater than the maximum value of the probable range, then only small improvements in the estimation of the model's parameters, their interactions, and/or the source term, should be necessary to provide reasonable assurance that a level of concern will not be exceeded. In other words, these minor improvements should result in the limiting value exceeding the maximum value of the probable range. The parameters that contribute most to the uncertainty in the model's prediction can be identified using appropriate techniques that measure the dependency of model predictions on parameter uncertainty (see accompanying paper on the analysis of parameter sensitivities submitted to SENES by Gardner and Hoffman).

iii) If the limiting value is clearly less than the maximum value, but greater than the minimum of the probable range, emphasis should be given toward making major improvements in the predictive accuracy of the model. Efforts should also be made to eliminate or reduce obvious sources of extreme conservatism that would be responsible for gross systematic error leading to large overpredictions. Attempts should be made to estimate parameters as a function of readily measureable site-specific variables. The possibility of covariance between parameters should be carefully considered. Site-specific investigations may be required to obtain more appropriate data. Expensive siting and source term modifications should be considered only after all attempts have been made to improve the estimates of the parameters and their interactions. On the other hand, siting and source term modifications that are clearly cost-effective (i.e. substantially above a level that would be considered marginal)

should be given serious consideration to comply with the principle of keeping radiological exposures "as low as is reasonably achieveable."

iv) If the limiting value is below the minimum value of the probable range, efforts should be directed toward reducing the source term or restricting the principal pathways of exposure. This latter step may require a change or modification of the site. Again, efforts should also be made to eliminate or reduce obvious sources of extreme conservatism that would be responsible for gross systematic errors leading to large overpredictions.

3.0 Measures of Dependency of Model Predictions On Parameter Uncertainty

When a limiting value is less than the maximum value of the probable range, there will be incentive to reexamine the effect of uncertain parameters on the distribution of predicted values. In so doing, it will be most effective to first focus attention on those parameters that contribute most to the uncertainty in the predicted values. As mentioned in our accompanying paper, as well as in numerous other publications, several methods exist for identifying these parameters.

The simplest method used to identify important parameters is that used for additive models. Assuming independence among parameters, the statistical variance determined for a given parameter is divided by the variance determined for the distribution of predicted values (Hoffman and Gardner 1983). For multiplicative models, the variance of the logarithms of parameter values can be divided by the variance of the log-transformed distribution of predicted values. These ratios are interpreted as the percent of the uncertainty in the model's prediction contributed by the parameter in question. The parameter with the highest ratio is the parameter with the highest importance. Additional terms are used to account for the covariance between parameters.

For more complex models whose parameter uncertainties have been propagated

with numerical methods, the measure of dependence between parameter and predictive uncertainty is best determined using statistical regression techniques. Since the errors propagated through these models are typically large (with coefficients of variation exceeding 100%) and since the response of the model's prediction is generally a monotonic and often nonlinear function of its parameters, simple linear regressions of input and output may give misleading results. Some type of transformation is usually necessary to produce a linear relationship between the model's parameters and its predictions. For models dominated by multiplicative processes, log-transformation may be useful. A more general method is to transform the values according to their rank order (from highest to lowest or vice versa). Rank transformation will force any monotonic relationship into one that is linear. This is the method we most frequently use to obtain reliable indices of parameter importance.

A practical indicator of parameter importance is the square of Spearman's Rank Correlation coefficient. This coefficient is obtained by correlating the rank transformed input and output through linear regression and squaring the result. This term represents the percent of the variance in the model's predictions accounted for by given parameter.

Another useful indicator is the relative partial of the sum of squares (Gardner, 1984). The strengths and weaknesses of many of these are discussed in the accompanying paper by Gardner and Hoffman on sensitivity analysis.

Once the indices of parameter importance have been selected and calculated, the analysis will identify the subset of the total number of parameters in the model that have a significant influence on the uncertainty in the model's prediction. Any improvement whatsoever in the estimation of these parameters and their interactions will produce a marked effect on the distribution of predicted values and the estimated 95% confidence interval "of sorts". Note that improved estimates of parameter uncertainties and interactions will usually have a more pronounced effect at the tails than at the midpoint of the distribution.

4.0 Summary Recommendations

i) Decisions based on the results of a parameter uncertainty analysis should refer to a calculated "probable range" of values rather than to extreme values in the tails, specific cumulative percentiles, or the exact shape of the distribution of predicted values.

ii) Decision-makers must recognize the potential influence of sources of systematic error on the results of a parameter uncertainty analysis.

iii) Estimates of parameter distributions should encompass the uncertainty associated with extrapolation to the specific assessment questions as well as that due to unexplained variability in the data.

iv) The probable range of predicted values can be conveniently based on a 95% confidence interval "of sorts", assuming the distribution of predicted values to be approximately lognormal.

v) If the uncertainty associated with the model's predictions is unacceptable (i.e. the maximum value of the calculated probable range exceeds some limiting value), a sensitivity analysis should be conducted to identify the most important model parameters. These are the parameters that should be selected for improved estimates of uncertainty and interaction.

vi) Parameter importance should be based on the dependence of the uncertainty in a model's prediction on the uncertainty within the estimate of a given parameter.

vii) A practical indicator of the dependence of the uncertainty in a model's prediction on the uncertainty in a given parameter is the square of Spearman's Rank Correlation coefficient. This index is efficient when used with numerical methods of simulating parameter error propogation. Other indices also exist that will produce similar results.

References

Aitchison, J., and Brown, J.A.C., 1969. "The Lognormal Distribution", Cambridge University Press, New Rochelle, N.Y.

Hoffman, F.O. and Gardner, R.H., 1983. "Evaluation of Uncertainties in Environmental Radiological Assessment Models." In: Radiological Assessments: A Textbook on Environmental Dose Assessment, J.E. Till and H.R. Meyer (eds.), U.S. Nuclear Regulatory Commission, Washington, D.C., NUREG/CR-3332, ORNL-5968.

Gardner, R.H., 1984. "A Unified Approach to Sensitivity and Uncertainty Analysis." Proceedings of the Tenth ISTED International Conference, Applied Simulation and Modelling, June 4-6, 1984, San Francisco, California.

Appendix E.4

Alternative Sampling Procedures: An Overview

Prepared by
Thomas Duever and Park Reilly
University of Waterloo, Dept. of Chemical Engineering

for
SENES Consultants Limited
July, 1984

1.0 Introduction

The objective of this survey is to determine what random sampling is, what the alternatives are and what the advantages and disadvantages associated with the various methods are. Furthermore a comparison of the major methods, simple random sampling, stratified random sampling and latin hypercube sampling for use in Monte Carlo modelling will be made. The information presented here represents opinions found in the literature. No conclusive answer was found as to which is the best method to be used for the required application. During the course of this survey, the application was kept in mind, that is the selection of values of input variables for Monte Carlo studies. As a result, sampling methods which are not applicable were not considered. In other words, a broader scope of references has been reviewed, but only the pertinent ones are discussed.

2.0 Alternative Sampling Methods

2.1 Simple Random Sampling (RS)

Raj (1968) gives the following definition of simple random sampling: "From a population of N units select one by giving equal probabilities to all units. Make a note of the unit selected and return it to the population. Perform

this operation n times to get a simple random sample of n units selected with replacement." Note that from the point of view of the random variable, this implies that each possible value is chosen according to its probability.

Random sampling gives an unbiased estimate of both the mean and variance. In addition, RS will asymptotically give the correct shape of the output distribution. Another advantage of RS lies in its simplicity.

In Figure 1, RS is illustrated for a situation where there are two input variables A and B. If a total sample of 16 sets is required, 16 independent observations are made on each of A and B. The sample consists of 16 pairs, each pair being made up of an observation on A and an observation on B.

2.2 Stratified Random Sampling (SRS)

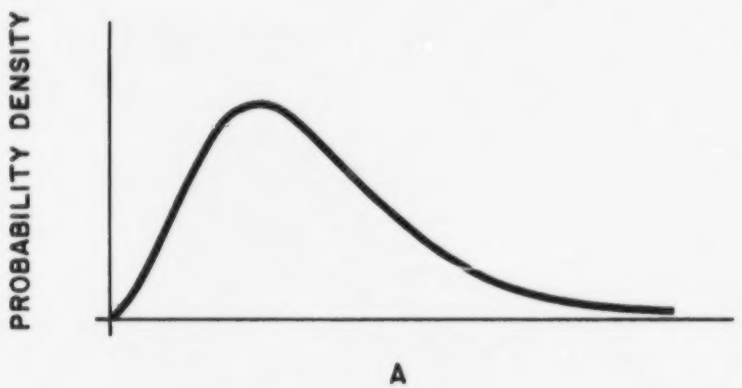
In SRS, all areas of the sample space S of $\underline{X}$ are represented by input values. The sample space S is partitioned into I disjointed strata, S_i, where $p_i = P(x \in S_i)$ represents the size of S_i. If I=1 this corresponds to RS. A random sample x_{ij}, $j=1,\ldots,n_i$ from S_i is then obtained such that $\sum_i n_i = N$.

This explanation of SRS given by McKay et al (1979) is somewhat deficient in that the values of I and of the n_i's are not specified. Presumably, $n_i = p_i N$ and $\Sigma n_i = N$.

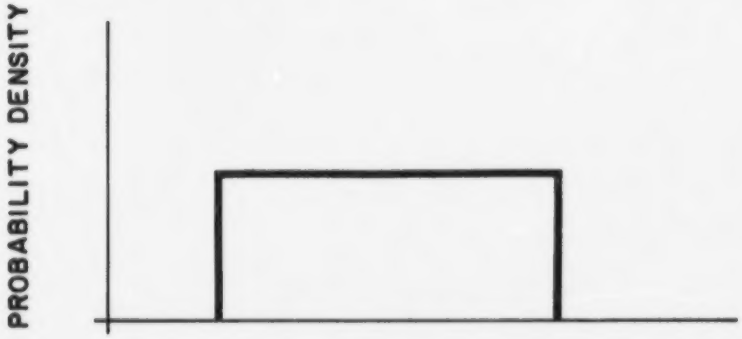
Consider the same example described for RS. The range of values of both A and B are divided into subranges by any derived method. Suppose the ranges of A and B are each divided into two parts at the median, as illustrated in Figure 2. Then if a total sample of 16 sets of values is required, four pairs of observations (a pair is one observation on each of A and B) are made by simple random sampling (RS) in each of the quadrants (strata) of Figure 2. Since the ranges of the variables are split at the median and since proportional stratification would likely (but not always) be used, each stratum has equal sampling in this case. Note that the ranges of A and B are not necessarily split at the centre. Here they are split at the median.

FIGURE 1

SIMPLE RANDOM SAMPLING (R.S.)



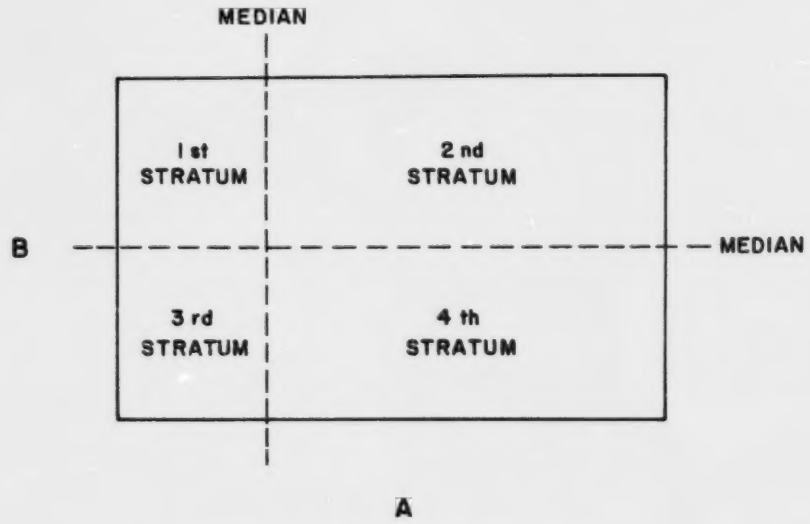
A



B

FIGURE 2

SYSTEMATIC RANDOM SAMPLING (S.R.S.)



Stratification may produce a gain in precision in the estimate of characteristics, such as the mean, of the whole population. It may be possible to divide a heterogeneous population into subpopulations which are internally homogeneous. If each stratum is homogeneous, precise estimates of any stratum mean can be obtained from a small sample. These precise estimates can then be combined to get a precise estimate for the whole population. In general then, one would expect smaller variances for estimators from SRS than from RS. This is not always true, in particular if the sample size for each stratum are far from the optimum (Cochrane 1977). The optimal stratum sample size is proportional to the product of the total number of units in the stratum and the variance within the stratum.

Jessen (1975) compares RS, simple stratification and two-way stratification or lattice sampling and concludes that lattice sampling gives the smallest variance and hence the most precise results.

2.3 Quota Sampling

In this approach to sampling, which has been widely used in opinion and market research surveys, the sample size required in each stratum is computed in advance so that stratification is proportional. Common variables for stratification in this type of problem are geographic area, age, sex and some measure of income level. The enumerator must then sample in a particular geographic area say, until the set quotas are fulfilled. If strict SRS were used, a great deal of sampling would be necessary to fulfill the quotas since in the latter stages, many samples chosen would fall into quotas already fulfilled. The filling of quotas is expedited by allowing for some nonrandom selection procedure. In other words, quota sampling can be viewed as stratified sampling with more or less nonrandom selection of units within strata. This can of course produce samples which are biased with respect to certain population characteristics.

2.4 Systematic Sampling (SYS)

This sampling method is quite different from simple random sampling. Cochrane (1977) gives the following explanation of systematic sampling. First the N units in the population are numbered 1 to N. Then a unit is chosen at random among the first k. Every kth unit is chosen thereafter to get a sample of size n.

The apparent advantages over simple random sampling are twofold:

1. It is easier to draw a sample.
2. It is intuitively more precise, about as precise as stratified random sampling with one sample unit per stratum with the units occurring at the same relative position. The systematic sample is more evenly spread out over the population and can therefore be considerably more precise than SRS.

A variant of systematic sampling involves selecting each unit at or near the center of each stratum. There are grounds which indicate that this is more precise than SYS initiated by choosing randomly among the first k units.

Systematic sampling can be considered to be a form of cluster sampling, since the choice of the unit at random among the first k is equivalent to choosing one of k clusters or groups out of the entire population. A systematic sample therefore is a simple random sample of one cluster unit from a population of k cluster units.

SYS is more precise than RS if the variance within the systematic samples is larger than the population variance as a whole. SYS is considered to be less precise than SRS. The performance of SYS is dependent upon properties of the population.

Another variation of SYS would be stratified systematic sampling. This is achieved by obtaining a systematic sample within each systematically chosen

stratum. This method will be more precise than SRS if systematic sampling within strata is more precise than simple random sampling within strata.

The disadvantages of SYS are that this method gives poor precision when unsuspected periodicity is present and that no trustworthy method for estimating the variance of the mean estimates from the sample data is known.

2.5 Latin Hypercube Sampling (LHS)

The last design is related to quota sampling, lattice sampling and to random balance designs discussed by Satterthwaite (1959). LHS takes the ideas of SRS one step further in that not only is each sample space of $\underline{X}$ represented, but each input variable X_k has all portions of its distribution sampled. To do so, the range of each input variable X_k is divided into N strata of equal marginal probability 1/N. Each stratum is sampled once. The components of the various X_k's are matched at random to give the N, k-dimensional input vectors X_i; i=1,...N. This method can be viewed as a k-dimensional extension of Latin-Square sampling.

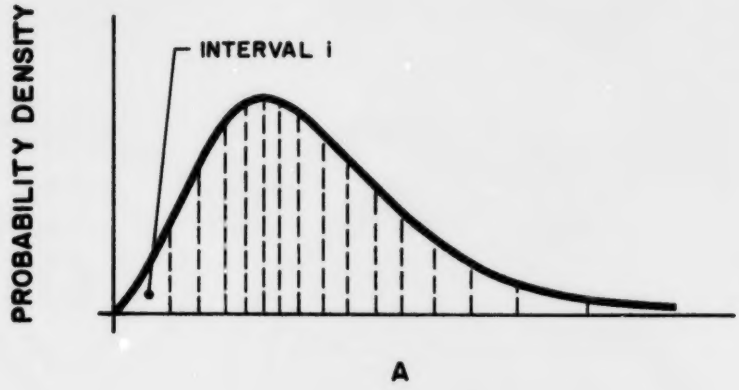
Consider once again the example used in Sections 2.1 and 2.2. To obtain a sample of size 16 on A and B, their ranges are broken up into 16 intervals of equal probability as shown in Figure 3. The intervals are then paired by a random process, for example by writing down the numbers from 1 to 16 and next to them a random permutation of the same numbers, for example:

A:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
B:	10	9	4	7	8	11	14	15	3	2	16	13	5	1	6	12

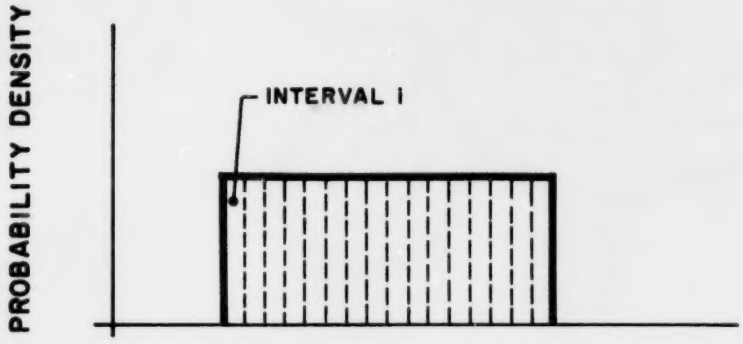
Then the sample is obtained by making a random observation in each interval of each pair by RS. As previously stated, the sample consists of 16 pairs of observations, one from A and one from B in each pair. The first pair would consist of an observation taken at random in the first interval of A along with one on the 10th interval of B, etc.

FIGURE 3

LATIN HYPERCUBE SAMPLING (L.H.S.)



A



B

NOTE:

VERTICAL LINES SHOWN ON PROBABILITY DENSITY FUNCTION ARE TO INDICATE EQUAL AREAS UNDER THE CURVE (N.T.S.)

The full design would have had 16^2 strata. LHS only uses the 1/16 fraction of these and therefore is a highly fractionated design.

SRS has the advantage over LHS in that the low and high value of each variable is combined probabilistically with the low and high value of every other variable. The advantage of LHS over SRS is that it ensures to a higher degree that all portions of the input variable distributions are sampled.

In summary, SRS is based on a complete factorial design whereas LHS is based on a highly fractionated factorial design.

It is interesting to note that while McKay et al (1979) divide each input into N strata of equal marginal probability, Helton and Iman (1982) divide the input into N intervals of either equal marginal probability or equal width. The latter is not advisable since is implies that all intervals are equally likely which they are not. They should be weighted according to their probability density. Subdivision on the basis of equal width will lead to a biased estimate of the mean and variance. In the "Program Users Guide," Iman et al (1980) suggest division according to probability only.

3.0 Comparison with Respect to Monte Carlo Calculations

In this section a comparison between random sampling (RS), stratified random sampling (SRS) and latin hypercube sampling (LHS) for selecting values of input variables, is made. The discussion is based on the papers by McKay et al (1979) and Iman and Conover (1980).

MacKay et al report the following results for a direct comparison of the three methods. First, all three methods give comparable, unbiased estimates of the sample means. The estimated standard deviation of the estimates show that $SD(\overline{Y}_{LHS}) < (SD(\overline{Y}_{SRS}) < SD(\overline{Y}_{RS})$ suggesting that LHS yields the most precise estimate. As pointed out by Easterling (Iman et al 1980), however, one serious drawback of LHS is that one cannot tell if LHS outperformed RS since

the variance of the estimator cannot be estimated from a single LHS run. The remedy is to run r LHS runs of size N/r rather than a single LHS run of size N. The latter is used by McKay et al to obtain the above comparison. In addition McKay et al determined that all three gave comparable estimates of the mean of the population variance. However, once again LHS resulted in a more precise estimate. Finally McKay et al report that all three methods yield comparable estimates of the mean of the cumulative distribution function (CDF) but, as before, LHS yields a better estimate. They conclude that LHS improves upon SRS which is better than RS. LHS improves upon RS when certain monotonicity conditions hold and it appears to be a good method to use for selecting values for input variables.

Lawrence (Iman et al 1980) points out that LHS must be used with caution since it may in fact give poorer estimates, in the sense of having larger variances, if monotonicity assumptions are not satisfied. Monotonicity here is meant with respect to the underlying function which relates $\underline{X}$ to Y. It must be continuous and monotone increasing or decreasing. This is of course hard to determine in advance since the function can be quite complicated.

Iman and Conover (1980) state that not all models allow for straightforward application of LHS, again citing the monotonicity conditions. In LHS the variance of the estimators is usually smaller than the variances of estimators obtained from other schemes; but this result may be closely related to the monotonicity property of the code. Population variance estimates do have some associated bias. However Iman and Conover claim that if the variance of estimators such as the mean is reduced this bias will be small.

Iman and Conover (1980) also claim that a very useful feature of LHS is that it allows one to measure sensitivity of the code to some of the assumptions of the model behind the code. In particular after a set of runs is completed under certain distribution assumptions on the input variables, the assumed distribution of the output variable may be made without making additional runs on the code. This claim is based on the assumption that each cell represents a small portion of the sample space in size and probability content since

there are N^k cells in total. Furthermore it is assumed that there is very little change in probability density over one sample cell and that hence the distribution can be assumed essentially uniform across the cell. Then it is assumed that different input distributions have very little effect on sampling as long as the different distributions have about the same range space S and are fairly smooth functions. This would be very useful since these codes often require a great deal of computer time to run. If the distribution function of the output variables is altered significantly by the adjustments to the input distribution then the output may be considered to be sensitive to assumptions regarding the input distribution.

McKay and Beckman (Iman and Conover 1980) suggest that this result should be used with "extreme caution: since it is not clear if errors made in misspecification of the input distribution will not add up to be significant when obtaining mean estimates, even though the effect of misspecification may be small for each specific cell".

Finally Easterling (Iman and Conover 1980) points out that as a design for sensitivity analysis, the latin hypercube is an example of the pure random balance design proposed by Satterthwaite (1959). It is an $N^{-(k-1)}$ fraction of an N^k design; where k is the number of variables; the fraction and confounding being selected at random. These designs when proposed were heavily criticised (see discussions following Satterthwaite (1959) by Youden, Kempthorne, Tukey, Box and Hunter). Easterling fails to see why such an inefficient design should be used for that particular type of problem.

4.0 Conclusions

Given the information just presented, only two sampling methods, simple random sampling (RS) and Latin Hypercube Sampling (LHS), should be considered for applications in probabilistic analysis. The former is easier to understand and to implement while the latter appears to give somewhat more consistent results in most cases. The final choice of sampling methods depends on the desired use. If the mean of output distributions under probabilistic

assessment is required, relatively few trials are implied and the larger computer storage requirements of LHS can be supported. However if the full output distribution is required, many more trials are needed and thus RS is likely to be preferred. Obviously there are many cases where there is no clear-cut preference between methods.

References

1. Cochran, W.G., 1977. "Sampling Techniques." John Wiley and Sons pub. New York.

2. Helton, J.C., and Iman, R.L., 1982. "Sensitivity Analysis of a Model for the Environmental Movement of Radionuclides." Health Physics, Vol. 42, No. 5, pp. 565-584.

3. Iman, R.L., and Conover, W.J., 1980. "Small Sample Sensitivity Analysis Techniques for Computer Models, with an Application to Risk Assessment." Communications in Statistics: Theory and Methods, Vol. A9, No. 17, pp. 1749-1842.

4. Iman, R.L., Davenport, J.M., and Zeigler, P.K., 1980. "Latin Hypercube Sampling (Program Users Guide)." Rep SAND79-1473, Sandia National LAboratories, Albuquerque, New Mexico.

5. Jessen, R.J., 1975. "Square and Cubic Lattice Sampling." Biometrics, Vol. 31, pp. 449-471.

6. McKay, M.D., Beckman, R.J. and Conover, W.J., 1979. " A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code." Technometrics, vol. 21, No. 2, pp. 239-245.

7. Raj, D., 1968. "Sampling Theory." McGraw Hill pub., New York.

8. Satterthwaite, F.E., 1959. "Random Balance Experimentation," Technometrics, Vol. 1, No. 2, pp. 111-137.

Appendix E-5

Software Quality Assurance (Q.A.): Current Concepts

Summary

This review attempts to summarize the vague, sometimes contradictory information available regarding software Quality Assurance (QA). Based on this review, some consistent definitions have been proposed as well as some practical suggestions for implementation on the probabilistic assessment project. The details contained in the body of the report are briefly summarized below.

For present purposes, software QA could be defined as "A planned and systematic pattern of all actions necessary to provide adequate confidence that the item or product conforms to established technical requirements" (IEEE/ANSI Standard 1981). A viable QA strategy should focus on error avoidance and error detection using an iterative verification and validation procedure. We define an error as being present "when the software does not do what the user reasonably expects it to do" (Myers 1976). Verification focuses on error avoidance by qualitatively examining the results of each step in the software development process. This can be done effectively using "structured walkthroughs" and is aided by the development of "data dictionaries".

The errors that inevitably remain after verification may be detected by selective testing during the validation phase. Test cases should be selected in some systematic manner to ensure maximum error detection and to increase product reliability. The key to effective validation in particular and QA in general, is a well documented, systematic approach.

1.0 INTRODUCTION

The reliability of computer software has been the object of much attention and expense since the advent of the computer age. The increasing use of computer

controls in critical systems (eg. nuclear power generation, strategic weapons systems) as well as notable software failures (eg. the first American space probe to Venus failed because of an error in a single Fortran statement (Myers 1976)). have led to increased emphasis on software 'Quality Assurance'. Despite this emphasis (and concurrent expense) a recent multiyear survey of over 200 projects indicated that "15% of all software projects never deliver anything and cost overruns of 100 to 200% are common. Software projects are sometimes considered successful when the overruns are held to 30% or when the user only junks a quarter of the results" (Bryan and Siegal 1984).

The current lack of adequate software QA stems, in part, from vague definitions and poor understanding of what QA is, or should be, about. By outlining current theoretical and practical QA techniques, some of the contradictions in the literature, as well as discrepancies between theory and practice, will become apparent. Using this variety of definitions, theories and common practices as a background, the following report attempts to outline a workable/acceptable QA strategy for the probabilistic assessment study.

1.1 Quality Assurance Defined

Many current definitions of QA are vague and general: i.e. "QA will include application of systematic programming techniques and the use of QA tests" (Lyon 1982) or "QA will ensure that a software product conforms to (meets or exceeds) prespecified standards" (Bryan and Siegal 1984). The Technical Advisory Committee (TAC) for the Canadian nuclear waste disposal program claimed that "Current state-of-the-art methods of assuring software reliability are already available at the operational level and are also very promising at the research level" (TAC 1981). Since the "tests" of the first definition and the "methods" of the third are unspecified, much latitude is allowed in selecting a QA methodology.

For example, the developers of the SYVAC2 program appear to have adopted a method known as "error seeding" as their final QA test (Andres 1982). This method involves deliberately placing a specified number of random, artificial

errors into the final program. By examining the relative number of "planted" and "natural" (unplanted) errors discovered by a testing team, some estimate of the number of natural errors remaining, and hence of the acceptability of the program can be made. In their own words, "The ultimate test of the quality of SYVAC2 in our approach will be (this) final QA test" (Andres 1982). However, a number of obvious flaws in this method were pointed out by J.D. Musa (1980). Musa claims that it is difficult in practice to generate random, artificial faults that are equivalent to natural faults. Since it is usually easier to find the seeded faults, the number of remaining natural faults will likely be underestimated using this method. There is also the problem of determining what is actually an "acceptable" number of natural faults remaining.

This example merely illustrates the problem of defining an acceptable QA method. For the sake of clarity, and with the goal of an "acceptable" definition in mind, it is probably best to use the definitions presented in a recent IEEE/ANSI Standard (1981). They define QA as **"A planned and systematic pattern of all actions necessary to provide adequate confidence that the item or product conforms to established technical requirements."** In this definition, it might be assumed that "adequate confidence" and "established technical requirements" imply mutual acceptability to both the program developer and ultimate user (ie. in this case the developer, SENES and the client, NUTP)

Since this definition of QA is still quite vague, it is important to examine what various authors consider to be adequate actions and technical requirements. In determining the reliability of a program, the terms verification and validation are often used. Since these terms generally refer to stages in the software development process, this process is outlined briefly below.

1.2 A Software Development Model

A macroscopic model of software development, as proposed by Myers (1976), is

presented in Figure 1. This model does not attempt to rigidly define stages that should occur in program development, but merely reflects the common 'top-down' programming structure (i.e. general to specific). The significant features of this model are the various translations that occur ('development activities' in Figure 1). These translations represent a communication process between the ultimate user and the software developer. Thus, software errors, representing communication breakdowns, mistakes and/or noise can arise during any one of these activities. A good QA strategy will attempt to avoid and detect faults during all of these processes. This is best accomplished by "verification and validation" of each developmental phase.

1.3 Verification and Validation Defined

As mentioned earlier, the terms verification and validation are encountered quite frequently in discussions of software QA. As with QA itself, the precise definitions are far from clear. Verification has been used to mean:

- testing a program in a simulated environment (Myers 1976)
- demonstrating (via inductive assertions and other mathematical proofs) the consistency between a computer program and specifications describing what the program is supposed to do (Yeh 1977)
- illustrating (via inspection, demonstration, analysis or test) that the software requirements are implemented into the design and code and that the code when executed meets the software requirements (IEEE/ANSI Standard 1981).

The most common definition, was recently expressed as "**the process of determining whether or not the products of a given phase of the software development cycle fulfill the requirements established during the previous phase**".(Boehm 1984). More succinctly, it implies constantly asking and answering the question, **"Am I building the product right?"**. This is probably the most useful definition, and as such, verification will form an integral part of the software QA strategy.

Validation is generally easier to define, although differing opinions do

Figure 1

MACRO MODEL OF SOFTWARE DEVELOPMENT

DEVELOPMENT ACTIVITY	**DEVELOPMENT SEQUENCE**
description of problem to be solved (statement of user requirements)	USER
translating requirements into program objectives	REQUIREMENTS
statement of external specifications (description of system behaviour from user's point of view)	OBJECTIVES
translate external specs into internal specs (flowcharting)	HIGH LEVEL DESIGN → DOCUMENTATION
coding	LOW LEVEL DESIGN
	PROGRAM CODE

exist. Some authors distinguish between validation and testing (Bryan and Siegal 1984), some claim that validation only occurs at the end of the software development process (Cooper and Fisher 1979), but most would agree that validation is the process of executing the software product and comparing test results to required performance. Validation, using this definition, usually involves examining the results of certain specified test cases.

1.4 A Software Quality Assurance Strategy

Thus, the two vital elements in a workable, acceptable software QA strategy will be an ongoing series of verifications and validations. This process will ensure "the highest possible quality, reliability and satisfaction of system needs and objectives" (Cooper and Fisher 1979) These two procedures, applied in an iterative manner will emphasize both error avoidance (verification) and error detection (validation) and thus help to ensure product reliability. To assist in applying this strategy to the task at hand, the concepts of verification and validation are discussed in more detail below.

Since "error" may be defined in several ways, a consistent definition of "error" is also needed. For present purposes it is reasonable to assume that an error is present "when the software does not do what the user reasonably expects it to do" (Myers 1976). Thus, an error can be present, even though the software performs according to specifications, if the specifications themselves are erroneous (see Figure 1).

2.0 VERIFICATION

Verification focuses on the early detection of errors and "getting the design right". Error detection and correction during design is far more cost-effective than error detection during testing (validation). "Software error cost studies from a large number of sources all confirm that it costs 20 to 30 times more to fix a requirement error during testing than if it were fixed during production of the requirements document" (Yeh 1977).

Verification requires the formalization of the top-down program design outlined in Figure 1. Although it costs time and money (and is likely to be tedious) to state explicitly in writing the objectives at each stage, experience has shown that overall this approach will save time and money. It is critical to ensure that both the user and developer agree on explicit requirements (general statement of goals) and on the detailed specifications, ie. what the program will do and how (in the user's terms) it will do it. Since the program developer will create his flowcharts, and ultimately his code, directly from these specifications, it is essential that they be understood correctly. Early analysis of the actual code also lends itself to the verification technique which follows.

Although verification of the coding of small program units can be done via mathematical techniques (inductive assertion, levels of abstraction, subgoal induction, predicate transformation, etc.) (Yeh 1977), that is not the goal of verification as previously defined. A technique that is useful for verifying most, if not all, of the program levels is known as the "structured walkthrough". This is simply an organized review of parts of software products by a group of diverse individuals. Generally the program developer guides two or three others (not intimately involved with the project) through his work. These objective participants attempt to find flaws in the product under review. A number of authors indicate that walkthroughs are a very effective procedure for enhancing product quality (Citron 1984, Andres 1982, Yourdon 1979).

Verification obviously relies on proper documentation of all the software development phases. Thus, the "data dictionaries" which define variables, the transfer of data between program units and the storage and retrieval of information are quite useful. They are excellent examples of how the formalization of intuitive "good programming" techniques is at the heart of software QA.

3.0 Validation

Although verification attempts to prevent errors in design and coding, such errors inevitably occur. Thus, the second phase of the QA strategy, validation, is required. Validation as mentioned earlier is often taken as synonomous with software testing (Myers 1976). The definition can be expanded, if desired, to include methods of predicting errors remaining in a program. These methods include the error seeding technique mentioned earlier, and may be useful as a final QA step. However, the main focus here, as in most software projects, will be on software testing (Murray 1984).

3.1 Software Testing

Again, before discussing the practicalities of software testing, it is useful to provide some background . Most authors emphasize that it is very difficult to prove that a large program is 100% correct. "**In general, it is impractical, often impossible to find all the errors in a program**" (Myers 1979). Since one cannot test a program to ensure that it is error free, and since it is fairly safe to assume that there are errors in a program, testing should concern itself with finding as many of these errors as possible. Testing, with this goal in mind, can help improve confidence that the program in fact does (sufficiently well) what it is supposed to do and more importantly does not do something that it is not supposed to do. The best way to establish confidence that the program is "perfect" is to try to disprove it; "that is, to try to find imperfections rather than just confirm that the program works correctly for some set of input data" (Myers 1979).

3.2 Test Case Selection

Given this functional definition of testing the next item which must be addressed is test case selection. That is "which subset of all possible test cases has the highest probability of detecting the most errors?" (Myers 1979). This remains an area of much debate and uncertainty as illustrated by statements such as: "Methods for finding reliable and valid test data

selection criteria are not yet well developed" (Yeh 1977) or "From an engineering point of view software testing has traditionally been ad hoc with few principles and no theoretical work indicating how to construct an adequate set of tests, or how to measure the adequacy of a set of test that someone else has constructed" (Goodenough and McGown, 1980). However, there does seem to be a consensus that a testing program must be ordered, or structured, in some way to be effective. Test cases must be selected based on certain criteria; a random set of test cases has little possibility of being an optimal set.

The selection of test cases depends on the program level being tested. Testing normally occurs on a program module (some small program unit or subunit - in the present case a submodel) as well as at higher levels. These higher levels often involve testing combinations of modules (integration testing) as well as testing of the entire system (system testing). The test case selection at each level requires an explicit set of test objectives and performance criteria for each program feature. A complete test case specification should specify the feature being tested, the input data and the expected output. Once the test cases are executed, the results should be compared with the expected results and also examined for other, unexpected results (eg. altered input data, etc). This structured testing is "in essence, a systematic and disciplined approach to carrying out the normal testing tasks in a manner that maximizes the effectiveness of time spent and minimizes the effort required for satisfactory test coverage" (Walsh 1977).

Test data have generally been chosen using two distinct methods. **Black box** testing involves test cases chosen from design specifications with no knowledge of how a program has been implemented. This method would require exhaustive testing over the entire input domain to detect all errors present. Conversely, **white box** test cases are selected with full knowledge of how a program has been implemented. This is generally done by an examination of the program logic, and thus may not detect specification errors. To be complete, white box tests must attempt exhaustive path testing, covering all possible logic paths through a program. Myers (1979) concludes that "Although

exhaustive input testing is superior to exhaustive path testing, neither prove to be useful strategies because both are infeasible". He advocates combining elements of both methods to achieve a systematic test strategy. In this way the black box methods will be used to determine whether a program satisfies its specifications while the white box methods will indicate internal errors in the program coding and logic.

Test cases at the module level should detect errors such as: missing control statements, inappropriate path selection or inappropriate or missing actions (DeMillo et al 1978). Of the many proposd black and white box testing methods it seems that an optimal(?) method combining elements from each category (as suggested earlier) can be developed to form an adequate test case selection strategy.

The most widely applicable and acceptable black-box testing method is known as "Boundary Value Analysis" (Myers 1979). It involves partitioning the input domain of a program into a set of "equivalence classes". That is, for each input condition a set of valid and invalid data is identified. Then, test cases are designed examining values just on, above and below the edge of each equivalence class. For example if an input parameter must be greater than 1.0, test cases where this parameter has the value of .99, 1.0, 1.01 should be examined. This can also be done if valid and invalid output ranges exist.

This boundary value test case selection criterion can be supplemented by test cases chosen using a familiar, intuitive, "error guessing" method. The program developer generally has an idea about what parts of a specification are particularly critical or error prone. These should be examined formally by constructing and evaluating appropriate test cases.

Since these black box methods may not excercise all parts of a program's logic flow, they should be supplemented by an appropriate white box method. To ensure that each decision in a program is executed and that each branch of a decision is exercised, "condition coverage" (Myers 1979) is a recommended method. This is only practical on smaller program units, but is an important

element in ensuring software quality. It involves supplementing the test cases designed during boundary value analysis with sufficient test cases to achieve the above goal. For programs that include multiple conditions (eg. IF (A=B) AND (C<D) DO 10) all possible combinations of these conditions should be examined. This is known as "multiple condition coverage" (Myers 1979).

The above strategy for module testing may seem too cumbersome and time consuming to be realistic. Again, it is a formalization of much of the intuitive testing process. Often in practise, the test cases are still selected on a rather ad hoc basis (Murray 1984), to achieve as complete coverage as a programmer decides. Metzger (1981) puts this informal approach in writing: **"I do not suggest imposing any rigid, formal module test scheme - only general guidelines. The programmer should put on paper, in his own words, and in his own format, the steps he proposes to execute in order to test the module".** Such a module test plan should then be subjected to a walkthrough. This may be a more realistic approach than the methods mentioned earlier, but it too should be done in some formal manner. A tentative format for test documentation is given in Figure 2.

At some point, module testing must cease and "higher order" or integration testing must begin. This involves testing groups of modules together and is largely a "black box" activity. Modules may be integrated, and thus this level of testing may proceed, in several ways. All modules may be tested separately and then integrated in one "big bang" near project completion; the testing may proceed in an incremental bottom-up or top-down manner (adding new untested modules to already tested ones); or, most efficiently, integration testing may proceed in a "sandwich" fashion (Myers 1976). In this method, testing proceeds both from the top down (i.e. general to specific) and bottom up (i.e. specific modules first) and hopefully testing is completed somewhere in the middle. Test case selection in this process may be quite arbitrary, but again, all test cases and results should be properly documented.

Figure 2

Tentative Test Case Documentation

Test Case #	Feature(s) Being Tested	Input Data Specifics	Expected Output	Test Output	Remarks

3.3 Test Completion Criteria

Given these testing methods, the final question to be addressed is "When to stop testing?" This is a difficult, if not impossible question to answer, since there is no way of knowing if the error just detected is the last remaining error (of potential significance). For module testing, if an approved set of test cases exist (selected arbitrarily by the programmer or in a more formal manner), testing should cease when these test cases all produce their prespecified, expected results. However, the test completion criteria for integration and system testing are far more vague and arbitrary. Some criteria which may be considered are: to test for a given time; to use the results of two independent test teams; to use the results of an error seeding test; or to prespecify the number of errors to be found (based on empirical averages) (Myers 1979). Another possible approach is to plot the errors found against time. Such a curve should reach an assymptote near "zero" (a suitably small number) errors found, as time progresses. Some point of little marginal return on the testing effort can be determined as testing proceeds, and used as the completion criterion.

To summarize the testing process inherent in software validation, some general guidelines can be highlighted (Myers 1979):

(i) A necessary part of a test case is a definition of the expected output or result.

(ii) A programmer should avoid testing his/her own program (this is a debatable point!)

(iii) The results of each test should be thoroughly inspected.

(iv) Test cases must be written for invalid and unexpected as well as valid and expected input conditions.

(v) Avoid throw-away (undocumented) test cases. Test cases represent an investment of time and effort and should therefore be documented and acknowledged

(vi) Do not plan a testing effort under the tacit assumption that no errors will be found.

(vii) The probablility of the existence of more errors in a section of a program is proportional to the number of errors already found in that section. This may seem counterintuitive but it has been empirically demonstrated.

3.4 An "Acceptance Test"

If the entire validation process, including test case selection, is subjected to walkthroughs and/or approval by the ultimate user, then the set of module and integration test cases can serve as a mutually agreeable "acceptance test". In this case, completion of the specified tests with no obvious errors will imply product acceptability.

4.0 Other Verification and Validation Techniques and Tools

For the sake of completeness, a brief list of several other tools and techniques which are available for software testing is provided. Some of these techniques are recently developed theoretical methods and thus, to the present time at least, are rarely used in practice, while others represent tools which have been used to test very large operating software systems (developed for instance by IBM or the U.S. military). They are only listed here; a number of more detailed references can be found in Myers (1979) (Chapter 8):

- module driver tools
- static flow analysis tools
- test coverage monitors
- mathematical proofs of program correctness
- program correctness provers
- symbolic execution
- test data generators
- environmental simulators
- sneak-circuit analysis
- virtual machines

- predictive models - Markov processes
- multiple linear regression analysis
- error seeding

5.0 CONCLUSIONS

In summary, a valid software QA program should be a well documented, iterative process of verification and validation. It is important to remember that "although testing has some bearing on software reliability, there is a limit to its value; reliability cannot be tested into a program" (Myers 1976). To obtain software that is reliable, robust and useful, "one must use appropriate QA techniques at each stage in the software development process" (Myers 1979) Given the vague, imprecise definitions and contradictions that exist in the literature, no generally acceptable/applicable QA strategy can be identified. A strategy applicable to a specific task should be designed based on the appropriate practical and theoretical considerations described earlier. Demillo et al (1978) sum up the notion that an appropriate QA strategy depends on the specific problem under consideration: **"We believe there is ample evidence in man's intellectual history that he does not solve important problems by viewing them from a distance. In fact, there is an Alice in Wonderland quality to fields which claim they can solve other peoples' problems without knowing anything in particular about the problems. So, there is certainly no need to apologize for applying ad hoc strategies in program testing. A programmer who considers his problems well, and skillfully applies appropriate techniques to their solution - regardless of where the techniques arise - will succeed".**

6.0 REFERENCES

Andres, T.H., 1982 "Rewriting SYVAC: Why and How". Environmental and Safety Branch, Whiteshell TR-2107.

Boehm, B.W., 1984. "Verifying and Validating Software Requirements and Design Specifications", IEEE Software Vol. 1, No. 1.

Bryan, W.L. and Siegal, S.G., 1984. "Product Assurance; Insurance Against a Software Disaster". Computer, Vol. 17, No. 4.

Citron, A., 1984. "A Software Review Method that Really Works". BYTE, No. 1.

Cooper, J.D. and Fisher, M.J. (eds), 1979. "Software Quality Management"; Petrocelli Books (NY).

DeMillo, R.A., Lipton, R.J. and Sayward, F.G., 1978. "Hints on Test Data Selection", IEEE Computer, Vol. 11, No. 4.

Druffel, E.D., 1983. "The STARS Program", IEEE Computer, Vol. 16, No. 11.

Foster, K.A., 1980. "Error Sensitive Test Case Analysis". IEEE Trans on Software Engineering, Vol. SE6, No. 3.

Goodenough, J.B. and McGown, C.L., 1980. "Software QA: Testing and Validation". Proceedings of IEEE, Vol. 68, No. 9.

IEEE/ANSI Standard 730-1981, 1981. "IEEE Standard for Software Quality Assurance Plans".

Leveson, N.G., 1984. "Software Safety in Computer Controlled Systems". Computer, Vol. 17, No. 2.

Lyon, R.B., 1982. "Nuclear Waste Disposal-Performance Assessment. Principles and Procedures". Canadian Nuclear Society. Proceedings of International Conference on Radioactive Waste Management.

Metzger, P.W., 1981. "Managing a Programming Project", Prentice Hall (NJ).

Murray, E.M., 1984. Software Specialist, DEC. Personal Communication (May, 1984).

Musa, J.D., 1980. "The Measurement and Management of Software Reliability". Proceedings of IEEE, Vol. 68, No. 9.

Myers, G.J., 1979 "The Art of Software Testing"; Wiley (NY).

Myers, G.J., 1976. "Software Reliability: Principles and Practices"; Wiley (NY).

Poston, R.M., 1984. "Software Standards", IEEE Software, Vol. 1, No. 1.

Redwine, S.T., 1983. "An Engineering Approach to Software Test Design". IEEE Trans. on Software Engineering, Vol. SE9, No.2.

SENES Consultants Limited, 1983. "Preliminary Probabilistic Assessment of the Long Term Effects of Uranium Mine Tailings in Canada." Proposal to the National Uranium Tailings Program Office.

Technical Advisory Committee (TAC) on the Nuclear Fuel Waste Management Program. 1980 - 1983.Annual Reports.

Walsh, D.A., 1977. "Structured Testing", Datamation, No. 7.

Yeh, R.T. (ed), 1977. "Current Trends in Programming Methodology: Vol II Program Validation"; Prentice Hall (NJ).

Yourdon, E., 1979. "Structured Walkthroughs"; Prentice Hall (NJ)

Zeil, S.J., 1983. "Testing for Perturbations of Program Statements". IEEE Trans. on Software Engineering, Vol. SE9, No. 3.

APPENDIX F

STUDY TEAM MEMBERS

APPENDIX F

STUDY TEAM MEMBERS

Dr. Douglas Chambers	SENES Consultants Limited
Dr. Leo Lowe	499 McNicoll Avenue
Dr. Jeno Scharer	Willowdale, Ontario M2H 2C9
Bruce Halbert	
Valerie Cassaday	
David Hopper	
Michael Murray	
W.C. Harrison	Scientific Authority National Uranium Tailings Program Canada Centre for Mineral and Energy Technology Energy, Mines & Resources Canada 555 Booth Street Ottawa, Ontario K1A 0G1
Dr. Richard Osborne	Atomic Energy of Canada Limited
Dr. Peter Barry	Chalk River Nuclear Laboratories
	Chalk River, Ontario K0J 1J0
Dr. Bruce Goodwin	Atomic Energy of Canada Limited
Dr. Terry Andres	Nuclear Fuel Waste Management
	Whiteshell Nuclear Research Establishment
	Pinawa, Manitoba R0E 1L0
Stan Feenstra	Golder Associates
Dr. Leo Rothenburg	3151 Wharton Way
	Mississauga, Ontario L4X 2B6

Dr. F. Owen Hoffman	121 Westwood Lane Oliver Springs, Tennessee U.S.A., 37840
Dr. Robert Gardner	107 Canterbury Road Oak Ridge, Tennessee U.S.A., 37830
Dr. Park Reilly	Reilly Consultants 48 Culpepper Driver Waterloo, Ontario N2L 5L1